

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:59:32 ; Search time 2512.03 Seconds  
(without alignments)  
8851.954 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373  
Sequence: 1 atgcattgttgaacttga.....gacaaatgttataaatg 1373

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_huv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	765.8	55.8	834	13	BI766766 603056866
2	753.2	54.9	948	14	BO707185 AGENCOURT
3	747.2	54.4	963	14	BO671259 AGENCOURT
4	743.4	54.1	777	13	BI819200 603034614
5	725.4	52.8	1071	14	BM921213 AGENCOURT
6	701.4	51.1	731	13	BI871711 603395825

7	688	50.1	828	13	BI596681	603243254
8	677.4	49.3	728 <td>13 <td>BI870393</td> <td>603395641</td> </td>	13 <td>BI870393</td> <td>603395641</td>	BI870393	603395641
9	584.4	42.6	609	13 <td>BI966060</td> <td>1e72904.x</td>	BI966060	1e72904.x
10	581.4	42.3	1027	17	AF163779	AF163779
11	562	40.9	940	14	BO884231	AGENCOURT
12	534.6	38.9	568	14	BM971606	UI-CF-EC1
13	531.6	38.7	910	12	BG110063	602279667
14	516.8	37.6	531	13	BI824443	603038693
15	493.6	36.0	918	12	BE577781	602092080
16	458.2	33.4	1033	11	AK020909	Mus muscu
17	454.4	33.1	561	10	AW676337	ur70409.y
18	444	32.3	951	14	BO674188	AGENCOURT
19	442.8	32.3	587	12	BG686319	602638232
20	440.2	32.1	456	13	BI966255	1e72904.y
21	436.8	31.8	440	13	BM128059	1e79b06.y
22	431.6	31.4	538	12	BF821434	MR1-RT003
23	404	29.4	436	17	AO890280	HS_3188-B
24	399.4	29.1	785	13	BI762908	603047966
25	394	28.7	413	9	AI422796	tE22e06.x
26	389.2	28.3	698	13	BI906850	603064633
27	374.2	27.3	416	9	AI231866	qna86c02.x
28	372.6	27.1	584	10	AW917574	EST348878
29	365.4	26.6	422	13	BI677255	1d87a02.x
30	365	26.6	399	9	AI913541	wa12c01.x
31	363.4	26.5	367	17	AO100365	HS_3054.A
32	359	26.1	474	13	BI965174	1d34e07.y
33	358.4	26.1	407	9	AI221985	q995a10.x
34	357.6	26.0	894	13	BI908274	603068526
35	352.8	25.7	409	10	AW131279	xFl61a04.x
36	347.2	25.3	397	9	AI669243	wc13g01.x
37	342.6	25.0	710	12	BE858778	7f95b06.x
38	341	24.8	377	13	BM509016	1h15b05.y
39	339.8	24.7	372	13	BI677256	1d87a02.y
40	330	24.0	374	13	BM505649	1h15b05.x
41	328	23.9	367	13	BM662622	UI-E-CQ1
42	327.8	23.9	650	12	BG404836	602420160
43	327	23.8	346	12	BE858822	7g38e04.x
44	327	23.8	471	9	AA221610	my18409.f
45	326	23.7	337	10	AM195034	xn45912.x

#### ALIGNMENTS

RESULT 1  
BI766766 834 bp mRNA linear EST 25-SEP-2001  
LOCUS 603056866F1 NIH\_MGC\_122 Homo sapiens CDNA IMAGE:5206217 5',  
DEFINITION mRNA sequence.  
ACCESSION BI766766  
VERSION BI766766.1 GI:15758344  
KEYWORDS  
SOURCE  
ORGANISM human.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LNA11517 row: C column: 18  
High quality sequence stop: 772.  
Location/Qualifiers

FEATURES  
source

1..834

[illegible]

OY	1049	-CCGCCACGGGCATCTCACAAGTCAAGGCCCCCATTCGACCTTGAGG	1102
Dd	780	CCCCCAAAGCCCATCTCACCTATTAGTCCCATTTCCCTGACTTTGGAGG	834
RESULT 2			
LOCUS	B0707185	948 bp	mRNA linear EST 16-JUL-2002
DEFINITION	AGENCOURT_8353983 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6278608		
ACCESSION	BQ707185		
VERSION	BQ707185.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 948)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@email.nih.gov Tissue Procurement: Dr. Mark Watson cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LILCM2466 row: n column: 17 High quality sequence start: 24 High quality sequence stop: 550. Location/Qualifiers  1..948 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6278608" /clone_id="NIH_MGC_113" /lab_host="DH10B (phage-resistant)" /note="Organ: spleen; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Nipalscript II RT (Life Technologies). Note: This is a NIH_MGC library."		
BASE COUNT	173 a 334 c 252 g 187 t	2 others	
ORIGIN			
Query Match	54.9%; Score 753.2; DB 14; Length 948;		
Best Local Similarity	95.6%; Pred. No. 4.7e-154;		
Matches	807; Conservative 0; Mismatches 30; Indels 7; Gaps 3;		
OY	305	ACCCTCGGAATGTAAATCCCAGACAGAAGAACGAGATCTCGGCTTCTCGAACC	364
Dd	1	ACCCTCGGAATGTAAATCCCAGACAGAAGAACGAGATCTCGGCTTCTCGAACC	60
OY	365	GACTAGTTGCGGCTCGCAGAAGTGCACCTAAAGGCCGGAANAACACGGGCTCGAAGAGCGA	424
Dd	61	GACTAGTTGCGGCTCGAAGAAGTGCACCTAAAGGCCGGAANAACACGGGCTCGAAGAGCGA	120
OY	425	TGCGAGCCCATTTATGAAGTTCAATCCAGACACTCGGACAGAGCGAGCGAGCGAGCTGTGG	484
Dd	121	TGCGAGCCCATTTATGAAGTTCAATCCAGACACTCGGACAGAGCGAGCGAGCGAGCTGTGG	176
OY	485	ACGGACAGTGAAGTGGCTGGGAGGAAGCAGAAATTAACAGCTCCAGCCCTGCGGCTTACA	544
Dd	177	ACGGACAGTGAAGTGGCTGGGAGGAAGCAGAAATTAACAGCTCCAGCCCTGCGGCTTACA	236
OY	545	ACCGCAATCGGGAGTTTTATGTCACCGGGGCTCTACTACTGTAAGTGTACAGG	604

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	237	ACGGCAGATGGGGAGGTTATAGTCACCCGGGCTCTACTACTGTACTGTACAG										
QY	605	TGCACCTTGATGAGGGAGAGGCTGTACCTGAGCTGACTTGCTGTGATGCTGGC										
Db	297	TGCATTTGATGAGGGAGAGGCTGTACCTGAGCTGACTTGCTGTGATGCTGGC										
QY	665	TGGCCCTGGGCTGGGAGGAATCTGAGCAGCTGGGGCAGTCCCTGGGGCCCCAGC										
Db	357	TGGCCCTGGGCTGGGAGGAATCTGAGCAGCTGGGGCAGTCCCTGGGGCCCCAGC										
QY	725	TCCGCTCTGGCAGGCTGTCTGGGCTGTTGGCCCTGAGGCCAGGGTCTCCCTGGCATCC										
Db	417	TCCGCTCTGGCAGGCTGTCTGGGCTGTTGGCCCTGAGGCCAGGGTCTCCCTGGCATCC										
QY	785	GCACCTCCCTGGGAGGCTCTGAGGCTGGCCCTCTCTCACCCTACTTGGAGCTTCC										
Db	477	GCACCTCCCTGGGAGGCTCTGAGGCTGGCCCTCTCTCACCCTACTTGGAGCTTCC										
QY	845	AGGTTCACTGAGGGGCCCTGTGCTCCGCCAGATGCTCCAGGCTCCGGCTCCCTGGAC										
Db	537	AGGTTCACTGAGGGGCCCTGTGCTCCGCCAGATGCTCCAGGCTCCGGCTCCCTGGAC										
QY	905	AGCTCTCTGGGAGCCCGGCTCCCTCTGCCCCACCTCAGCGCTTCTTGGTCTCAGA										
Db	597	AGCTCTCTGGGAGCCCGGCTCCCTCTGCCCCACCTCAGCGCTTCTTGGTCTCAGA										
QY	964	GCCCTCTCTGAGAGGCTGCTGGGCTGTTTCACTGTTTCAATCCACATTAATACA										
Db	657	GCCCTCTCTGAGAGGCTGCTGGGCTGTTTCACTGTTTCAATCCACATTAATACA										
QY	1024	GTAATCCCACTCTTAATCTTACACTCCGCCACCGGCCACTCTCACTCACTAGCTCCC										
Db	717	GTAATCCCACTCTTAATCTTACACTCCGCCACCGGCCACTCTCACTCACTAGCTCCC										
QY	1084	AATCCCTGAGCC -TTTGAGGCCGCCAGATGCTGATCCCTCCCTGGGCCAGAGCCCC										
Db	777	AATCCCTGAGCCCTTTGAGGCCGCCAGAGGATCTAACCCTCCCTGGGCCAGAGCCCC										
QY	1142	AGGG 1145										
Db	837	AAGG 840										
RESULT 3												
LOCUS	B0671259											
DEFINITION	AGNCOURT_8303564 NIH_MGC_102 Homo sapiens CDNA clone IMAGE:6274716											
ACCESSION	B0671259											
VERSION	B0671259.1											
KEYWORDS	EST.											
SOURCE	human.											
ORGANISM	Homo sapiens											
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.											
TITLE	1 (bases 1 to 963)											
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/											
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)											
	Unpublished (1999)											
	Contact: Robert Strussberg, Ph.D.											
	Email: cgaqbs-remail.nih.gov											
	Tissue Procurement: ANCC											
	CDNA Library Preparation: Rubin Laboratory											
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)											
	DNA Sequencing by: Agencourt Bioscience Corporation											
	Clone distribution: MGC clone distribution information can be											

Query Match	Best Local Similarity	Score	DB	Length
Matches 868; Conservative	91.9%;	747.2;	DB 14;	963;
			Pred. No. 9.5e-153;	
			Mismatches 63;	Indels 13; Gaps 7;
<pre> /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6274716" /clone_11b="NIH_MGC_102" /tissue_type="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: salivary gland; Vector: pORF7; site_1: XhoI; Site-2: EcoRI. cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: This is a NIH_MGC library." </pre>				
BASE COUNT	194 a	326 c	260 g	183 t
ORIGIN				
Query Match	54.4%;	Score 747.2;	DB 14;	Length 963;
Best Local Similarity	91.9%;	Pred. No. 9.5e-153;		
Matches 868; Conservative	0;	Mismatches 63;	Indels 13;	Gaps 7;
369	AGTTGGGCTCCGACAGATGTCACCTTAAGGCGGAAACAGCGGCTCCAGAGCGATCGC	428		
1	AGTTGGGCTCCGACAGATGTCACCTTAAGGCGGAAACAGCGGCTCCAGAGCGATCGC	59		
429	AGCCCATTTATGAAGTTCAATCCAGACCTTGACAGGACGAGGCGAGGCGAGGTGTGACGG	488		
60	AGCCCATTTATGAAGTTCAATCCAGACCTTGACAGGACGAGGCGAGGCGAGGTGTGACGG	119		
489	GACAGTAGTGGCTGGGAGGAAGCCAGAAATCAACAGCTTCAGCCCTTGCGGTACACCG	548		
120	GACAGTAGTGGCTGGGAGGAAGCCAGAAATCAACAGCTTCAGCCCTTGCGGTACACCG	179		
549	CCAGATCGGGGAGGTTATATAGTACACCGGCGTGGCTCTACTACCTGATGTCAGGTGA	608		
180	CCAGATCGGGGAGGTTATATAGTACACCGGCGTGGCTCTACTACCTGATGTCAGGTGA	239		
609	CTTTGATGAGGGGAGGAGCTGTACTCTGAAGTGGACTGTCTGTGGATGATGTGCTGGC	668		
240	CTTTGATGAGGGGAGGAGCTGTACTCTGAAGTGGACTGTCTGTGGATGATGTGCTGGC	299		
669	CCTGGGCTCGTGAAGAAATTTCTACGCCATCGCGCACTTCCCTGGGCGCCAGTCCG	728		
300	CCTGGGCTCGTGAAGAAATTTCTACGCCATCGCGCACTTCCCTGGGCGCCAGTCCG	359		
729	CCCTGCGCAGGTGTGGGCTGTGGGCGCTGGGCGCAGGAGTCTCCCTGGATCGGAC	788		
360	CCCTGCGCAGGTGTGGGCTGTGGGCGCTGGGCGCAGGAGTCTCCCTGGATCGGAC	419		
789	CCCTCCCTGAGGCCATCTCAAGGCTGCCCTCTCTCACTACTTGTGAGCTCTTCAGGT	848		
420	CCCTCCCTGAGGCCATCTCAAGGCTGCCCTCTCTCACTACTTGTGAGCTCTTCAGGT	479		
849	TCACTAGAGGGGCTGTGTCTCCACACAGTGTCTCCACAGGCTGCCGCTCCCTGACAGCT	908		
480	TCACTAGAGGGGCTGTGTCTCCACACAGTGTCTCCACAGGCTGCCGCTCCCTGACAGCT	539		
909	CTCTGGGACCGGCGTCCCTCTGCGCCACCTCAGCGGCTCTTTGGTCCAGACCTGGCCC	968		
540	CTCTGGGACCGGCGTCCCTCTGCGCCACCTCAGCGGCTCTTTGGTCCAGACCTGGCCC	599		
969	TCCTCTTAGAGCTGCTGGGCTGTTCACGTTTTCATCCATCCCATTAATACAGTATT	1028		
600	TCCTCTTAGAGCTGCTGGGCTGTTCACGTTTTCATCCATCCCATTAATACAGTATT	659		
1029	CCCACTTTATTTACAACTCCCC- ACCGCCACTCTTCACCTCACTAGTCCGCAATC	1087		
660	CCCACTTTATTTACAACTCCCC- ACCGCCACTCTTCACCTCACTAGTCCGCAATC	719		
1088	CC- ---TGACCTTTGAGGCGCCCAAGTATCTGCACTCCCGCTGGCCACATACCCCGAG	1143		
720	CCCTGACCCCTTTGAAGGCGCCCAAGGATCTGCACTCCCGCTGGCCACATACCCCGCA	779		

QY	1144	GGCATT-CTGTTCACTGACATCTGTGGGCAAGG-ATGGGCTCCAGAAAGCCCACTTCAGG	1201
Db	780	GGCATTGGGGCTTACTGACAACTCTGGGGGGGAGGAAGAGGGGTCCAAAAGACCCCTTTTGGG	839
QY	1202	CACTAAGAGGGGCT--GGACCTGGGGCGGACGAAGCCAAAGAGACTGGG---CCTAGGCCA	1256
Db	840	ACTTAAAGGGGCTTGGAGACTGGGGGGGAGAAAGAACCCAAAAAATGTGGGGCCCAAGGCCCA	899
QY	1257	GGAGTTCCTCCAAATGTGAGGGCGGAGAAACAGACACACTCTCC	1300
Db	900	GAAATTCCTCCAAAGGGAGGGGGCGGAAAAAACAAAAAGCTCC	943
RESULT 4			
LOCUS	B1B19200		
DEFINITION	B1B19200	777 bp	mRNA
ACCESSION	603034614F1	NIH_MGC_115	Homo sapiens cDNA clone IMAGE:5175698 5', mRNA sequence.
VERSION	B1B19200		
KEYWORDS	B1B19200.1	GI:15930750	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC htcp://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Life technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LNL11437 Row: 1 Column: 03 High quality sequence stop: 759.		
FEATURES			
source	1..777		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="5175698"		
	/clone_id="NIH_MGC_115"		
	/lab_host="DH10B"		
	/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Grueter (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC library."		
BASE COUNT	140 a	268 c	215 g
ORIGIN			154 t
Query Match	54.1%	Score 743.4	DB 13; Length 777;
Best Local Similarity	99.1%	Pred. No. 6.2e-152;	
Matches 758; Conservative	0;	Mismatches 6;	Indels 1;
		Gaps	1.1;
QY	288	GGCAGAGGAGGACGACGAGCCGTCGACACTGATCCCGACAGAGAAAGCCAGATCC	347
Db	13	GGCAGAGGAGGACGACGAGCCGTCGACACTGATCCCGACAGAGAAAGCCAGATCC	72
QY	348	TGCGGCTTTTCTGGAACGCACTAGTTCGGGCTCGCAAGAGTGACCTTAAAGGCCGGAAC	407
Db	73	TGCGGCTTTTCTGGAACGCACTAGTTCGGGCTCGCAAGAGTGACCTTAAAGGCCGGAAC	132
QY	408	ACGGGCTCGAAGAGCGATGCGACGCCCATTTATGAAGTTTCACGACGACTGACAGAGCG	467

Db	133	ACGGGCTCGAAGACGATCGACGCCATTATGAGTTTCATCCAGCACTGGACAGGACGG	192
QY	468	AGGCCAGGACAGGTGTGACGGGACAGTGAATGGCTGGGAGAGAGCCAGATCAACAGCTC	527
Db	193	AGGCCAGGACAGGTGTGACGGGACAGTGAATGGCTGGGAGAGAGCCAGATCAACAGCTC	252
QY	528	CAGCCCTCTGGGCTACACAACCCGACAGATCCGGGAGATTATATCAACCCGGGCTGGGCTCTA	587
Db	253	CAGCCCTCTGGGCTACACAACCCGACAGATCCGGGAGATTATATCAACCCGGGCTGGGCTCTA	312
QY	588	CTACCTGTACGTGATAGGAGCACTTTGATGAGGGGAGAGCTCTCTACCTGAAGCTGACTT	647
Db	313	CTACCTGTACGTGATAGGAGCACTTTGATGAGGGGAGAGCTCTCTACCTGAAGCTGACTT	372
QY	648	GCTGGTGGATGGTGTGCTGAGCCCTGCGCTGCGCTGGAGAGAAATTCAGGCACTGCGGCGAG	707
Db	373	GCTGGTGGATGGTGTGCTGAGCCCTGCGCTGCGCTGGAGAGAAATTCAGGCACTGCGGCGAG	432
QY	708	TTCCCTCGGGCCCCCAGCTCCGCTCTGCGCAGGTGTCTGGGCTGTTGGCCCTCGGGCCAGG	767
Db	433	TTCCCTCGGGCCCCCAGCTCCGCTCTGCGCAGGTGTCTGGGCTGTTGGCCCTCGGGCCAGG	492
QY	768	GTCCTCCCTGGGGAATCCGCGCACCCCTCCCTGGGAGCCATCTCAAGGCTGCGCCCTTCTCTAC	827
Db	493	GTCCTCCCTGGGGAATCCGCGCACCCCTCCCTGGGAGCCATCTCAAGGCTGCGCCCTTCTCTAC	552
QY	828	CTACTCGGACTCTTCCAGGTTTCACTGAGGGGACCCCTGGTCTCCGCCACAGTGGTCCAGGC	887
Db	553	CTACTCGGACTCTTCCAGGTTTCACTGAGGGGACCCCTGGTCTCCGCCACAGTGGTCCAGGC	612
QY	888	TGCGGCGTCCCTCGACAGCTCTCTGGGCAACCCGCTCCCTCTGCGCCACCTCAAGCCG	947
Db	613	TGCGGCGTCCCTCGACAGCTCTCTGGGCAACCCGCTCCCTCTGCGCCACCTCAAGCCG	672
QY	948	TCTTTGTCACGACCTGCGCCCTCCCTCTAGAGGGCTGCGTGGGACCTGTCAGGTGTTTCC	1007
Db	673	TCTTTGTCACGACCTGCGCCCTCCCTCTAGAGGGCTGCGTGGGACCTGTCAGGTGTTTCC	732
QY	1008	ATCCACATTAATACAGTATTCGCCACTCTTAATCTTACACTCCG	1052
Db	733	ATCCACATTAATACAGTATTCGCCACTCTTAATCTTACACTCCG	776
RESULT 5			
LOCUS	BM921213	1071 bp	mRNA linear EST 12-MAR-2002
DEFINITION	AGENCOURT 6633046 NIH_MGC_115 Homo sapiens cdna clone IMAGE:5752561		
ACCESSION	BM921213		
VERSION	BM921213.1	GI:19371592	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 1071)		
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:gsapbs-remail.nih.gov">gsapbs-remail.nih.gov</a>		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	Plate: LLAM12786 row: p column: 02		
	High quality sequence stop: 656.		
FEATURES	Location/Qualifiers		
SOURCE	1..1071		



```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752561"
/clone_1ib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT      226 a 346 c 279 g 220 t
ORIGIN

Query Match      52.8%; Score 725.4; DB 14; Length 1071;
Best Local Similarity 93.4%; Pred. No. 5.6e-148;
Matches 836; Conservative 0; Mismatches 46; Indels 13; Gaps 7;

OY 412 GCTGGAAGAGCGATCGACGCCATTATGAGTTCATCCAGACCTGACAGAGGAGCG 471
    |||||||
Db 12 GCTGGAAGAGCGATCGACGCCATTATGAGTTCATCCAGACCTGACAGAGGAGCG 71

OY 472 CAGGAGGTGGAGCGGAGTGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 531
    |||||||
Db 72 CAGGAGGTGGAGCGGAGGAGTGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 131

OY 532 CCTGTGGGCTACAAACCGGACAGATCGGGAGTATATGACACCGGGCTGTACTATAC 591
    |||||||
Db 132 CCTGTGGGCTACAAACCGGACAGATCGGGAGTATATGACACCGGGCTGTACTATAC 191

OY 592 CTGTACTGTCAAGGTGCACTTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 651
    |||||||
Db 192 CTGTACTGTCAAGGTGCACTTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 251

OY 652 GTGATGTGTGTGGGCGGCTGCGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 711
    |||||||
Db 252 GTGATGTGTGTGGGCGGCTGCGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 311

OY 712 CTCGGGGCCCAAGCTCCGCTCTGCGAGGTGTGTGGGCTGTGGGCTGTGGGCTG 771
    |||||||
Db 312 CTCGGGGCCCAAGCTCCGCTCTGCGAGGTGTGTGGGCTGTGGGCTGTGGGCTG 371

OY 772 TCCTGTGGGAGTCCGACCTCTCCCTGGGCGCATCTCAAGGCTGCCCTCTCTAC 831
    |||||||
Db 372 TCCTGTGGGAGTCCGACCTCTCCCTGGGCGCATCTCAAGGCTGCCCTCTCTAC 431

OY 832 TTGGGACCTTCCAGGTCTCAAGTGAAGGGGCGTGTCTCCCAAGTCTCCAGGCTGCC 891
    |||||||
Db 432 TTGGGACCTTCCAGGTCTCAAGTGAAGGGGCGTGTCTCCCAAGTCTCCAGGCTGCC 491

OY 892 GGTCTCCCTGACAGCTCTGTGGGACCGGTCCTCTGCGCCCACTTCAAGCGCTCTT 951
    |||||||
Db 492 GGTCTCCCTGACAGCTCTGTGGGACCGGTCCTCTGCGCCCACTTCAAGCGCTCTT 551

OY 952 TGTCTCAGACCTGCGCTCTCTAGAGAGGTGCTGGGCTGTTCACGTGTTTCCATCC 1011
    |||||||
Db 552 TGTCTCAGACCTGCGCTCTCTAGAGAGGTGCTGGGCTGTTCACGTGTTTCCATCC 611

OY 1012 CACATTAATAGTATTTCCCTCTTATCTTACAACTCCCGACCGCCCACTCTCCACT 1071
    |||||||
Db 612 CACATTAATAGTATTTCCCTCTTATCTTACAACTCCCGACCGCCCACTCTCCACT 671

OY 1072 CACTAGCTCCCAATCCCTGTGAGCCCTTTGAGGCGCCCACTGATCTGCACT -CCGCCCTGGC 1130
    |||||||
Db 672 CACTAGCTCCCAATCCCTGTGAGCCCTTTGAGGCGCCCACTGATCTGCACTCCGCCCTGGC 731

OY 1131 CACAGAGCCCG -AGGGAATGTG -TTCACTGTACTCTGTGGGCAAGGA -TGGGTCCAGA 1186
    |||||||
Db 732 CACAGAGCCCGCAAGGCAATGTGTGTTTCACTGTGAGCTGTGGGCAAGGAATGGGTCCAGA 791

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405478"
/clone_1ib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pcmv-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      125 a 242 c 228 g 136 t
ORIGIN

Query Match      51.1%; Score 701.4; DB 13; Length 731;
Best Local Similarity 99.3%; Pred. No. 9e-143;
Matches 726; Conservative 0; Mismatches 1; Indels 4; Gaps 2;

OY 1187 AGACCCCACTTCAGG----CACTAAGAGAGGCTGGAGCTGGCGGAGG----AGCCAAAG 1239
    |||||||
Db 792 AAACCCCACTTCAGGCACTTAAAGAGGCGCTGGAGCTGGCGGAGGAGGAGGAGGAGG 851

OY 1240 AGACTGGG-CCTAGGCGCAGAGTTCCTCAATGTGAGGGGCGAGAAACAGACAG 1293
    |||||||
Db 852 AAACCTGGGCGCTAGGCGCAGAGGTCCTCAATGTGGGAGGAGGCGCAGAAACCAAG 906

RESULT 6
BI871711 731 bp mRNA linear EST 11-OCT-2001
LOCUS 603395825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5',
DEFINITION mRNA sequence.
ACCESSION BI871711 GI:16045366
VERSION BI871711.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 731)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LHAM12034 row: b column: 07
High quality sequence stop: 728.
Location/Qualifiers
1. 731
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405478"
/clone_1ib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pcmv-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      125 a 242 c 228 g 136 t
ORIGIN

Query Match      51.1%; Score 701.4; DB 13; Length 731;
Best Local Similarity 99.3%; Pred. No. 9e-143;
Matches 726; Conservative 0; Mismatches 1; Indels 4; Gaps 2;

OY 132 CCTGGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 251
    |||||||
Db 1 CCTGGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60

OY 252 GCTGTCCGCGC--CAGGAGCGTGGCCAGGAGAGTGTGGGAGAGAGGAGGAGGAGG 308
    |||||||
Db 61 GCTGTCCGCGCAGGAGGAGTGTGGGAGAGGAGTGTGGGAGAGGAGGAGGAGGAGG 120

OY 309 GTGGAAGTGAATCCCGAGACAGAAAGAGAGATCTGCGCTTCTGGAACGACT 368
    |||||||
Db 121 GTGGAAGTGAATCCCGAGACAGAAAGAGAGATCTGCGCTTCTGGAACGACT 180

OY 369 AGTTGGGCTTGAGAGAGTGCACCTTAAAGCGCGGAAACAGAGGCTCGAAGAGCATGC 428
    |||||||
Db 181 AGTTGGGCTTGAGAGAGTGCACCTTAAAGCGCGGAAACAGAGGCTCGAAGAGCATGC 240

```

```

Oy 429 AGCCATATGAAATTATCATCAGACCTGAGACGAGAGGAGCGAGAGTGTGACGG 488
    |||||||
Db 241 AGGCCATTTATGAAGTTATCATCAGACCTGAGACGAGAGGAGCGAGAGTGTGACGG 300
Oy 489 GACAGTGAATGGCTGGAGAGAGCCAGAAATCAACAGCTCCAGCCCTGTGGCTTACAACG 548
    |||||||
Db 301 GACAGTGAATGGCTGGAGAGAGCCAGAAATCAACAGCTCCAGCCCTGTGGCTTACAACG 360
Oy 549 CCAGATCGGGAGATTATATGATCACCAGGCGTGGCTCTACTACTCTAGTCAAGTGC 608
    |||||||
Db 361 CCAGATCGGGAGATTATATGATCACCAGGCGTGGCTCTACTACTCTAGTCAAGTGC 420
Oy 609 CTTTGTATGAGGGAGAGCTGTCTACTGAACTGAACTGAACTGAACTGAACTGAACTG 668
    |||||||
Db 421 CTTTGTATGAGGGAGAGCTGTCTACTGAACTGAACTGAACTGAACTGAACTGAACTG 480
Oy 669 CCGGCGCTGCTGGAGAAATTCACAGCACTGCGGCGAGTCCCTGGGGCCCGACGTCGG 728
    |||||||
Db 481 CCGGCGCTGCTGGAGAAATTCACAGCACTGCGGCGAGTCCCTGGGGCCCGACGTCGG 540
Oy 729 CCGTCCGAGGTGTCTGGGCTGTGGGCTGTGGGCGGAGGCTCCCTGCGGATCCGAC 788
    |||||||
Db 541 CCGTCCGAGGTGTCTGGGCTGTGGGCTGTGGGCGGAGGCTCCCTGCGGATCCGAC 599
Oy 789 CCGTCCGAGGTGTCTGGGCTGTGGGCTGTGGGCGGAGGCTCCCTGCGGATCCGAC 848
    |||||||
Db 600 CCGTCCGAGGTGTCTGGGCTGTGGGCTGTGGGCGGAGGCTCCCTGCGGATCCGAC 659
Oy 849 TCACCTAGAGGGGCTGTGGGCTGTGGGCGGAGGCTCCCTGCGGATCCGAC 908
    |||||||
Db 660 TCACCTAGAGGGGCTGTGGGCTGTGGGCGGAGGCTCCCTGCGGATCCGAC 719
Oy 909 CTCCTGGGACAC 919
    |||||||
Db 720 CTCCTGGGACAC 730

RESULT 7
B1596681 828 bp mRNA linear EST 07-SEP-2001
LOCUS 603243254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',
DEFINITION B1596681
ACCESSION B1596681
VERSION B1596681
KEYWORDS B1596681 GI:15489620
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 828)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloning Strategy: In-cyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1M11722 row: k column: 13
High quality sequence stop: 776.
Location/Qualifiers
1. 828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285892"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"

```

```

/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT 155 a 278 c 223 g 172 t
ORIGIN
Query Match 50.1%; Score 688; DB 13; Length 828;
Best Local Similarity 95.3%; Pred. No. 7.8e-140;
Matches 732; Conservative 0; Mismatches 30; Indels 6; Gaps 2;
Oy 312 GGAACATGATCCCAACAGAAAGAGAGTCTGGGCTTCTGACCGACTAGT 371
    |||||||
Db 43 GGAACATGATCCCAACAGAAAGAGAGTCTGGGCTTCTGACCGACTAGT 102
Oy 372 TCGGCTCGGAGAGTGCACCTTAAAGCGGAAACAGGCGTCAAGAGCATCGAC 431
    |||||||
Db 103 TCGGCTCGGAGAGTGCACCTTAAAGCGGAAACAGGCGTCAAGAGCATCGAC 162
Oy 432 CCATTATGAAGTTATCATCAGACCTGAGACAGGAGCGAGGCTGTGACGGAC 491
    |||||||
Db 163 CCATTATGAAGTTATCATCAGACCTGAGACAGGAGCGAGGCTGTGACGGAC 222
Oy 492 AGTGAATGGCTGGAGAGAGCCAGATCAACAGCTCAGCCCTCGGCTACAAACGCCA 551
    |||||||
Db 223 ATTGAATGGCTGGAGAGAGCCAGATCAACAGCTCAGCCCTCGGCTACAAACGCCA 282
Oy 552 GATCGGGAGATTATATGATCACCAGGCTGGGCTCTACTACCTGATGACGTCGACT 611
    |||||||
Db 283 GATCGGGAGATTATATGATCACCAGGCTGGGCTCTACTACCTGATGACGTCGACT 342
Oy 612 TGATGAGGGAGAGGCTGTCTACTGAACTGAGCTGTGTGTGTGTGTGTGTGTGTGT 671
    |||||||
Db 343 TGATGAGGGAGAGGCTGTCTACTGAACTGAGCTGTGTGTGTGTGTGTGTGTGTGT 402
Oy 672 GCGCTGCTGAGAGAAATTCACAGCACTGCGGCACTGCGGCACTGCGGCACTGCGGCT 731
    |||||||
Db 403 GCGCTGCTGAGAGAAATTCACAGCACTGCGGCACTGCGGCACTGCGGCACTGCGGCT 462
Oy 732 CTGCGAGGTGTCTGGGCTGTGGGCTGTGGGCGGAGGCTCTCCCTGCGGATCCGAC 791
    |||||||
Db 463 CTGCGAGGTGTCTGGGCTGTGGGCTGTGGGCGGAGGCTCTCCCTGCGGATCCGAC 522
Oy 792 CCCCTGGGCGCATCTCAAGGCTGCCCTTCTCAGCTTCTGAGCTTCTGAGGTTCA 851
    |||||||
Db 523 CCCCTGGGCGCATCTCAAGGCTGCCCTTCTCAGCTTCTGAGCTTCTGAGGTTCA 582
Oy 852 CTGAGGGGCGCTGTGCCACAGTGTGCCAGGCTGTGCCAGGCTGTGCCAGGCTGTGC 911
    |||||||
Db 583 CTGAGGGGCGCTGTGCCACAGTGTGCCAGGCTGTGCCAGGCTGTGCCAGGCTGTGC 642
Oy 912 TGGGACACCGGCTCCCTGTGCCACAGTGTGCCAGGCTGTGCCAGGCTGTGCCCTCC 971
    |||||||
Db 643 TGGGACACCGGCTCCCTGTGCCACAGTGTGCCAGGCTGTGCCAGGCTGTGCCCTCC 702
Oy 972 CTTTGAAGGCTGCTGGGCTGTTCACGTTGT----TTCCATCCCAATTAATACAGT-- 1025
    |||||||
Db 703 CTTTGAAGGCTGCTGGGCTGTTCACGTTGTATATATATATATATATATATATATAT 762
Oy 1026 ATTCCACCTTTATCTTAAACATCCCGGACGCGGCTGCAACGCTCA 1073
    |||||||
Db 763 TTTCCACCTTTATCTTAAACATTTCCCGGACGCGGCTGCAACGCTCA 810

```

```

RESULT 8
B1870393 728 bp mRNA linear EST 11-OCT-2001
LOCUS B1870393
DEFINITION 603395641F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405459 5',

```

ACCESSION	mRNA sequence.
VERSION	B1870393
KEYWORDS	B1870393.1 GI:16044066
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.
COMMENT	1 (bases 1 to 728) National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strusberg, Ph.D. Email: cgephs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LHAM12034 Row: a Column: 12 High quality sequence stop: 728. Location/Qualifiers 1..728 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_1lib="IMAGE:5405459" /clone_1lb="NIH_MGC_90" /tissue_type="adenocarcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: This is a NIH_MGC library."
FEATURES	
source	
BASE COUNT	125 a 240 c 227 g 136 t
ORIGIN	
Query Match	49.3%; Score 677.4; DB 13; Length 728;
Best Local Similarity	99.0%; Pred. No. 1,6e-137;
Matches 724; Conservative	0; Mismatches 1; Indels 6; Gaps
OY	192 CCTGGCGCTGAGCCTTCGCTCGGCCCTGCTGGCCCTGTCACTTTGGGAGCGCGGCATC 251
Dd	1 CCTGGCGCTGAGCCTTCGCTCGGCCCTGCTGGCCCTGTCACTTTGGGAGCGCGGCATC 60
OY	252 GCTGTCGCC---CAGGAGCCTGCCAGAGAGAGCTGTGTGGCAGAGAGACCAGAGACC 308
Dd	61 GCTGTCGCCCCAGCAGAGAGCCTGCCAGAGAGAGCTGTGTGGCAGAGAGACCAGAGACC 120
OY	309 GTGCGAAGTGAAATCCCCAGACAGAGAAGAAAAGCCAGAGATCCTGGCCCTTCTGAACGACT 368
Dd	121 GTGCGAAGTGAAATCCCCAGACAGAGAAGAAAAGCCAGAGATCCTGGCCCTTCTGAACGACT 180
OY	369 AGTTGGGCTCGCAGAGTAGCACCTAAAGCGCGGAAAAACAACGAGGCTCGAAGAGCGATTGC 428
Dd	181 AGTTGGGCTCGCAGAGTAGCACCTAAAGCGCGGAAAAACAACGAGGCTCGAAGAGCGATTGC 240
OY	429 AGCCCATTTATGAAGTTTCATCCAGACCTTGGACAGAGAGGCGGAGGCGAGGTTGTGACCG 488
Dd	241 AGCCCATTTATGAAGTTTCATCCAGACCTTGGACAGAGAGGCGGAGGCGAGGTTGTGACCG 300
OY	489 GACACTGATGAGTCTGGGAGAGAGCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCG 548
Dd	301 GACACTGATGAGTCTGGGAGAGAGCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCG 360
OY	549 CCAGATGCGGGAGATTATAGTACCCGGGCTGGGCTCTACTACTCTGTAAGTGTGAGGTGCA 608
Dd	361 CCAGATGCGGGAGATTATAGTACCCGGGCTGGGCTCTACTACTCTGTAAGTGTGAGGTGCA 420
OY	609 CTTTATATAGGGGAGGAGGCTGTCTACTGAAAGCTGAACTTGTGTGTGTGTGTGCTGGC 668
Dd	420 CTTTATATAGGGGAGGAGGCTGTCTACTGAAAGCTGAACTTGTGTGTGTGTGTGCTGGC 668

Db	Accession	Version	Keywords	Source	Organism	Title	Journal	Comment
Db	421	CTTTATGATGAGGGAAGGCTGTC	TACTAGAGTGAAGTCTGCTGGTGGATGTCGCTGC	480				
Oy	669	CCTGGCTGCTCC	TGAGAGGAATTC	TACGACCACTGCGGCGACGTTCCCTCGGGCCCCAGCTCCG	728			
Db	481	CCTGGCTGCTCC	TGAGAGGAATTC	TACGACCACTGCGGCGACGTTCCCTCGGGCCCCAGCTCCG	540			
Oy	729	CCTGCGCAGAGGTC	TGAGGCTGTTGGCCCTGCGGCGACAGGTC	CTCCGAGATCCGAC	788			
Db	541	CCTGCGCAGAGGTC	TGAGGCTGTTGGCCCTGCGGCGACAGGTC	CTCCGAGATCCGAC	599			
Oy	789	CCTCCCTGAGGCGCAT	CTCAAGAGCGTCCGCTCCACCTACTTTCGACACTTC	TCCAGAGT	848			
Db	600	CCTCCCTGAGGCGCAT	CTCAAGAGCGTCCGCTCCACCTACTTTCGACACTTC	TCCAGAGT	658			
Oy	849	TCACCTGAGGCGCGCTG	GTCTCCGACAGTGTCTCCAGGCTGCGGCGCTCCCTCGACAGCT	908				
Db	659	TCACCTGAGGCGCGCGCTG	GTCTCCGACAGTGTCTCCAGGCTGCGGCGCTCCCTCGACAGCT	717				
Oy	909	CTCTGCGGACCG	919					
Db	718	CTCTGCGGACCG	919					
RESULT 9								
BI966060/c								
LOCUS	BI966060	609 bp	mRNA	linear	EST 12-MAR-2002			
DEFINITION	1e/2944.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens							
FEATURES	cdna clone IMAGE:5672623 3' similar to TR:054907 054907 TNF-RELATED							
ACCESSION	BI966060							
VERSION	BI966060.1	GI:16340465						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.							
AUTHORS	1 (bases 1 to 609) Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blustain, A., Schmitt, A., Theising, B., Rittner, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarelis, R., Williams, T., Jackson, Y., and Bowers, Y.							
COMMENT	Endocrine Pancreas Consortium Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Unpublished (2000) Endocrine Pancreas Consortium Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Harvard University, Howard Hughes Medical Institute Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138 Tel: 617-495-1812 Fax: 617-495-8557 Email: dmelton@oebio.harvard.edu Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fuchs.harvard.edu) This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov High quality sequence stop: 412. Location/Qualifiers 1..609 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5672623" /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1" /sex="Both" /tissue_type="Islets of Langerhans" /dev_stage="Adult" /lab_host="DH10B" /note="Organ: Pancreas; Vector: pSPORT1; Site:1; Not 1; Site:2; Sal 1; Starting library constructed using Superscript Plasmid Library Kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column							

fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:793-806. 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 142 a 128 c 209 g 129 t 1 others  
ORIGIN

Query Match 42.6%; Score 584.4; DB 13; Length 609;  
Best Local Similarity 99.7%; Pred. No. 3.1e-117;

Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 ACCCTCCCTGGGCGCATCTCAAGGCTGCGCCCTCTCTACTACTCTGGAGCTCTCCAG 846  
DB 609 ACCCTCCCTGGGCGCATCTCAAGGCTGCGCCCTCTCTACTACTCTGGAGCTCTCCAG 550  
QY 847 GTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 906  
DB 549 GTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 490  
QY 907 GTTCTGAGGACCGGCTGCTGCGGCGTCCACAGGCTGCTGCTGCTGCTGCTGCTGCTG 966  
DB 489 GTTCTGAGGACCGGCTGCTGCGGCGTCCACAGGCTGCTGCTGCTGCTGCTGCTGCTG 430  
QY 967 CCTCCCTAGAGGCTGCTGCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026  
DB 429 CCTCCCTAGAGGCTGCTGCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 370  
QY 1027 TTCCCACTCTTACTTACAACTCCCGCACCGGCTCTCACTCTCACTAGCTCCCAAT 1086  
DB 369 TTCCCACTCTTACTTACAACTCCCGCACCGGCTCTCACTCTCACTAGCTCCCAAT 310  
QY 1087 CCTGACCTTTTGAAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGAGCCCGCAGGGC 1146  
DB 309 CCTGACCTTTTGAAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGAGCCCGCAGGGC 250  
QY 1147 ATTTGTTTACTGTACTGTGTGGGCAAGGATGGTTCAGAAAGCCCACTTACAGGACTA 1206  
DB 249 ATTTGTTTACTGTACTGTGTGGGCAAGGATGGTTCAGAAAGCCCACTTACAGGACTA 190  
QY 1207 AGAGGGGCTGAGCTGTGCGGCGAGGAAGCAAGAGACTGGGCTAGAGGCAAGATTCCCA 1266  
DB 189 AGAGGGGCTGAGCTGTGCGGCGAGGAAGCAAGAGACTGGGCTAGAGGCAAGATTCCCA 130  
QY 1267 AATGTGAGGGGCGAGAAACAAGCAAGCTCTCTGAGAAATTCCTGTGATTTTAA 1326  
DB 129 AATGTGAGGGGCGAGAAACAAGCAAGCTCTCTGAGAAATTCCTGTGATTTTAA 70  
QY 1327 AACAGATATTATTATTATTATTGTGCAAAATGTTGATAAATG 1373  
DB 69 AACAGATATTATTATTATTATTGTGCAAAATGTTGATAAATG 23

RESULT 10  
AF163779 1027 bp DNA linear GSS 29-AUG-2000  
LOCUS AF163779 Human Homo sapiens genomic clone BAC750E14, DNA sequence.  
DEFINITION AF163779  
ACCESSION AF163779  
VERSION AF163779.1 GI:5726439  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
JOURNAL Cousin, P., Billotte, J., Chabert, P. and Shaw, P. H.  
Physical map of 17p13 and the genes adjacent to p53  
Genomics 63 (1), 60-68 (2000)

MEDLINE 20130114  
COMMENT Contact: Shaw PH  
Experimental Oncology  
Institute of Pathology  
Rue du Bugnon 25, Lausanne, VD 1011, Switzerland  
sub.clone=AB28 Asc-BamHI PSL1180  
Class: BAC subclone.

FEATURES  
source  
1..1027  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="17p"  
/clone="BAC750E14"  
/note="Vector: pMO5 Blue"

BASE COUNT 207 a 317 c 282 g 208 t 13 others  
ORIGIN

Query Match 42.3%; Score 581.4; DB 17; Length 1027;  
Best Local Similarity 98.0%; Pred. No. 1.5e-116;

Matches 577; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 785 GCACCTCCCGGGGCGCATCTCAAGGCTGCGCCCTCTCTACTACTCTGGAGCTCTCCAG 844  
DB 1 GCACCTCCCGGGGCGCATCTCAAGGCTGCGCCCTCTCTACTACTCTGGAGCTCTCCAG 60  
QY 845 AGTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 904  
DB 61 AGTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 120  
QY 905 AGCTCTGTGGGCAACCGGCTGCTGCGGCGTCCACAGTGTCCAGGCTGCTGCTGCTGCTG 964  
DB 121 AGCTCTGTGGGCAACCGGCTGCTGCGGCGTCCACAGTGTCCAGGCTGCTGCTGCTGCTG 180  
QY 965 CCCCTCCCTAGAGGCTGCTGCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1024  
DB 181 CCCCTCCCTAGAGGCTGCTGCGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 1025 TATTCACACTTACTTACAACTCCCGCACCGGCTCTCACTCTCACTAGCTCCCA 1084  
DB 241 TATTCACACTTACTTACAACTCCCGCACCGGCTCTCACTCTCACTAGCTCCCA 300  
QY 1085 ATCCCTGACCTTTTGAAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGAGCCCGCAGG 1144  
DB 301 ATCCCTGACCTTTTGAAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGAGCCCGCAGG 360  
QY 1145 GCATTGTGTTCACTGTACTGTGCGGCGAGGATGGTTCAGAAAGCCCACTTACAGGAC 1204  
DB 361 GCATTGTGTTCACTGTACTGTGCGGCGAGGATGGTTCAGAAAGCCCACTTACAGGAC 420  
QY 1205 TAAGAGGGGCTGAGCTGTGCGGCGAGGAAGCAAGAGACTGGGCTAGAGGCAAGATTCC 1264  
DB 421 TAAGAGGGGCTGAGCTGTGCGGCGAGGAAGCAAGAGACTGGGCTAGAGGCAAGATTCC 480  
QY 1265 CAAATGTGAGGGGCGAGAAACAAGCAAGCTCTCTGAGAAATTCCTGTGATTTT 1324  
DB 481 CAAATGTGAGGGGCGAGAAACAAGCAAGCTCTCTGAGAAATTCCTGTGATTTT 540  
QY 1325 AAAACGATATTATTATTATTATTATTGTGCAAAATGTTGATAAATG 1373  
DB 541 AAAACGATATTATTATTATTATTATTGTGCAAAATGTTGATAAATG 589

RESULT 11  
BO884231 940 bp mRNA linear EST 16-AUG-2002  
LOCUS BO884231  
DEFINITION AGENCOURT\_8682031 lupskl\_sciatic\_nerve Homo sapiens cDNA clone  
IMAGE:6197488 5', mRNA sequence.  
ACCESSION BO884231  
VERSION BO884231.1 GI:22276239  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 940)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs@mail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
 Plate: L1AM13607 row: 3 column: 17  
 High quality sequence stop: 453.

FEATURES  
 source  
 Location/Qualifiers  
 1..940  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6197488"  
 /clone\_lib="Lupski-sciatic.nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; CDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCG-3' and 5'-GACTAGTCTAGTACGGAGCGCGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 170 a 325 c 255 g 190 t  
 ORIGIN

Query Match 40.9%; Score 562; DB 14; Length 940;  
 Best Local Similarity 96.5%; Pred. No. 2,6e-112;  
 Matches 660; Conservative 0; Mismatches 15; Indels 9; Gaps 8;

QY 301 CAGGACCCGTCGGAATCCACAGAGAGAAAGCCAGATCTGGCCCTTCTG 360  
 DB 1 CAGAGCCCGTCGGAATCCACAGAGAGAAAGCCAGATCTGGCCCTTCTG 60  
 QY 361 AACGACTAGTTCGGCTCGCAGAGAGTCCACTAAAGCCGGAACACAGGGCTCGAGA 420  
 DB 61 AACGACTAGTTCGGCTCGCAGAGAGTCCACTAAAGCCGGAACACAGGGCTCGAGA 120  
 QY 421 GCGATCGGAGCCATTAGAGTTCATCCAGACTCGACAGACAGAGAGAGAGAGAGAGT 480  
 DB 121 GCGATCGGAGCCATTAGAGTTCATCCAGACTCGACAGAGAGAGAGAGAGAGAGAGT 180  
 QY 481 GTGACGGGACAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 540  
 DB 181 GTGACGGGACAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 240  
 QY 541 TACAACGGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTCTACTACCTGTACTGT 600  
 DB 241 TACAACGGCCAGATCGGGAGTTTATAGTCAACCGGGCTGGGCTCTACTACCTGTACTGT 300  
 QY 601 CAGGTGACCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 660  
 DB 301 CAGGTGACCTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 360  
 QY 661 GTGCTGGCCCTGCGCTGCTGGAGAGATTCACAGCACTGCGGCGAGTCCCTCGGGGCC 720  
 DB 361 GTGCTGGCCCTGCGCTGCTGGAGAGATTCACAGCACTGCGGCGAGTCCCTCGGGGCC 420  
 QY 721 CAGCTCGCCCTCTG-CCAGGTGTCTGGGCTGTGGCCCTCGGCGCA-GGGTCTCCCTG 778

DB 421 CAGCTCGCCCTCTG-CCAGGTGTCTGGGCTGTGGCCCTCGGCGCA-GGGGCTCCCTG 480  
 QY 779 GATCCGACACCTCCCTGGGCGCAATCTCAGGCTG-CCCTCTCTCCTACCTAC-TCGG 836  
 DB 481 GGATCCGACACCTCCCTGGGCGCAATCTCAGGCTG-CCCTCTCTCCTACCTACCTG 540  
 QY 837 ACTCTTCAGGTTCACTAGAGGGGCGCTGTCCCCACAGTGT-CCAGGCTGGCGG- 894  
 DB 541 ACTCTTCAGGTTCACTAGAGGGGCGCTGTCCCCACAGTGT-CCAGGCTGGCGGCT 600  
 QY 895 TCCCTCGACAGCTCTCT-GGGACACCGGCTCCCTGTGGCCAGCC--TCACCCGCTCTT 951  
 DB 601 TCCCTCGACAGCTCTCTGTGGAGAACCGGCTCCCTGTGGCCAGCCCTCAGCGGCTT 660  
 QY 952 TGCTCCAGACCTGCCCCCTCCCTCT 975  
 DB 661 TGCTCCAGACCTGCCCCCTCCCTCT 684

RESULT 12  
 BM971606/c 568 bp mRNA linear EST 21-MAR-2002  
 LOCUS BM971606  
 DEFINITION UI-CF-EC1-ab1-p-06-0-UI.s1 UI-CF-EC1 Homo sapiens CDNA clone  
 ACCESSION UI-CF-EC1-ab1-p-06-0-UI.3', mRNA sequence.  
 VERSION BM971606  
 KEYWORDS BM971606.1 GI:19589193  
 EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 568)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171

Email: paul.mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com)  
 The following repetitive elements were found in this CDNA sequence: 1-82, >AT-rich/low-complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=yes.

FEATURES  
 source  
 Location/Qualifiers

1..568  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EC1-ab1-p-06-0-UI"  
 /clone\_lib="UI-CF-EC1"  
 /tissue\_type="Lung"  
 /dev\_stage="Adult and Fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-CF-EC1 is a normalized CDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded CDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is

AGGCTTAC  
TAG-LIB-UI-CF-ECL  
TAG-TISSUE-Normal Lung Epithelial Cells Tissue nos 369-371  
and 380-383  
TAG-SEQ-NAAGTCTTAC"

BASE COUNT 130 a 119 c 184 g 135 t

JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM10019 row: e column: 18  
High quality sequence strop: 493.

FEATURES  
source location/Qualifiers  
1. 910  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4367225"  
/clone\_lib="NIH MGC 86"  
/tissue\_type="osteosarcoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: bone; Vector: PCMV-SPORT6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.533 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."

Query Match 38.9%; Score 534.6; DB 14; Length 568;  
Best Local Similarity 99.3%; Pred. No. 2.3e-106;  
Matches 537; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

BASE COUNT 262 a 279 c 225 g 144 t

ORIGIN

Query Match 38.7%; Score 531.6; DB 12; Length 910;  
Best Local Similarity 88.6%; Pred. No. 1.1e-105;  
Matches 622; Conservative 0; Mismatches 74; Indels 6; Gaps 4;

QY 833 TCGGACTCTTCAGGTTCTGAGGGGCTGTGCTCCGACAGTGTCCAGGCTGGCG 892  
DB 568 TCGGACTCTTCAGGTTCTGAGGGGCTGTGCTCCGACAGTGTCCAGGCTGGCG 509  
QY 893 GCTCCCTTCAGAGCTCTCTGGGCAACCGGCTCTGCTCCGACGCTGCTTT 952  
DB 508 GTTCCCTTCAGAGCTCTCTGGGCAACCGGCTCTGCTCCGACGCTGCTTT 449  
QY 953 GCTCCAGAGCTCTCTCTGAGAGCTGCTGGGCTGTTACAGTGTTCATGCC 1012  
DB 448 GCTCCAGAGCTCTCTCTGAGAGCTGCTGGGCTGTTACAGTGTTCATGCC 389  
QY 1013 ACATAAATACATATTTCCACATCTTATCTTACACCTCCGACGCTCTCCACCTC 1072  
DB 388 ACATAAATACATATTTCCACATCTTATCTTACACCTCCGACGCTCTCCACCTC 329  
QY 1073 ACTAGCTCCCAATCCCTGACCTTTGAGGCGCCAGTATCTGACCTCCGCTGGCA 1132  
DB 328 ACTAGCTCCCAATCCCTGACCTTTGAGGCGCCAGTATCTGACCTCCGCTGGCA 269  
QY 1133 CAGACCCCGAGGATGTTGTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1192  
DB 268 CAGACCCCGAGGATGTTGTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 209  
QY 1193 CACTTCAGGCTAAGAGGGGGCTGGGCTGGGCGGAGGAAAGCAAGACATGGGCTTGA 1252  
DB 208 CACTTCAGGCTAAGAGGGGGCTGGGCTGGGCGGAGGAAAGCAAGACATGGGCTTGA 149  
QY 1253 GCCAGAGTGTCCCAATGTGAGGGGCGAGAAACAAGACATGCTCTGAGATTCG 1312  
DB 148 GCCAGAGTGTCCCAATGTGAGGGGCGAGAAACAAGACATGCTCTGAGATTCG 89  
QY 1313 CTGTGGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1372  
DB 88 CTGTGGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 29  
QY 1373 G 1373  
DB 28 G 28

RESULT 13  
BG110063 910 bp mRNA linear EST 30-JAN-2001  
LOCUS 60227967F1 NIH\_MGC\_86 Homo sapiens cDNA clone IMAGE:4367225 5',  
DEFINITION mRNA sequence.  
ACCESSION BG110063  
VERSION BG110063.1 GI:12603569  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 910)  
AUTHORS NIH-MGC http://mgi.ncl.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

QY 627 TGTCTACTGAAAGCTGAGCTGCTGCTGATGTTGTTGCTGCTGCTGCTGAGGA 686  
DB 1 TGTCTACTGAAAGCTGAGCTGCTGCTGATGTTGTTGCTGCTGCTGCTGAGGA 60  
QY 687 ATTCTAGCCACTGCGGCGCAATTCCTGCGGCGCCAGCTCGGCTGTGCTGCTGCTG 746  
DB 61 ATTCTAGCCACTGCGGCGCAATTCCTGCGGCGCCAGCTCGGCTGTGCTGCTGCTG 120  
QY 747 GGTGTGGGCTGCGGCGCAAGGCTGCTGCTGCGGATCGGACCTGCTGCGGCGCATCT 806  
DB 121 GGTGTGGGCTGCGGCGCAAGGCTGCTGCTGCGGATCGGACCTGCTGCGGCGCATCT 179  
QY 807 CAAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866  
DB 180 CAAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239  
QY 867 CTCCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926  
DB 240 CTCCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299  
QY 927 CTCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986  
DB 300 CTCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359  
QY 987 GGGGCTGTTCAAGTGTTCATCCCATTAATACATTAATACATTAATACATTAATACATTA 1046  
DB 360 GGGGCTGTTCAAGTGTTCATCCCATTAATACATTAATACATTAATACATTAATACATTA 419  
QY 1047 CTCCGACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106  
DB 420 CTCCGACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478  
QY 1107 CAGTATCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166  
DB 479 AGTATCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538  
QY 1167 TGGGCAAGGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226  
DB 539 GAGGCAAGGATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595  
QY 1227 CAGGAAGCCAAAGAGACTG-GGCTTGGGCGGAGCTTCCCAATGTGAGGGGCGAGAAAC 1285

Db	596	AGGAGACGAAGAAGAGACTGAGGCGCTTAGGCAAGCAGCAATTTCCAAATGACGGGCGAGAAAA	655
QY	1286	AAGACAGCTCCCTCCCTTGAGAAATTCCTGTGGATTTTAAA	1327
Db	656	CAAGGAAGACCTCCCATGAGATGCGCTGAGAGATATGACAA	697
RESULT 14			
LOCUS	Bi824443	531 bp	mRNA linear EST 04-OCT-2001
DEFINITION	60303699F1 NIH_MGC_115	human sapiens cDNA clone IMAGE:5179510 5'	
ACCESSION	Bi824443		
VERSION	Bi824443.1	GI:15935993	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.		
TITLE	1 (bases 1 to 531)		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov. Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN). DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LHAM11447 row: j column: 23 High quality sequence stop: 529.		
FEATURES			
Source	Location/Qualifiers		
	1..531		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5179510"		
	/clone_11b="NIH_MGC_115"		
	/lab_host="DH10B"		
	/note="Organ: Pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."		
BASE COUNT	72 a 223 c 119 g 117 t		
ORIGIN			
Query Match	37.6%;	Score 516.8;	DB 13; Length 531;
Best Local Similarity	99.4%;	Pred. No. 1.7e-102;	
Matches 529;	Conservative 0;	Mismatches 2;	Indels 1;
		Gaps 1;	
QY	661	GTGCTGGCCCTGCGCGCTTGAGGAATTCACAGCAGTGGGCGCAGTCCCTCGGGGCC	720
Db	1	GTGCTGGCCCTGCGCGCTTGAGGAATTCACAGCAGTGGGCGGCGATTCCTCGGGGCC	60
QY	721	CAGCTCCGCGCTTGCGCAGGTGTCTGGGCGTGTGGCCCTGCGGCGAAGGCTCTCCCGCGG	780
Db	61	CAGCTCCGCGCTTGCGCAGGTGTCTGGGCGTGTGGCCCTGCGGCGA-gtctctccctg	119
QY	781	ATCCGACCCCTCCCTGGGCGCATCTCAAGCTGCCCCCTTCCTCAGCTTGGGATC	840
Db	120	ATCCGACCCCTCCCTGGGCGCATCTCAAGCTGCCCCCTTCCTCAGCTTGGGATC	179
QY	841	TTTCAGGTTCACTGAGGGCCCTGGTCTCTCCCAAGTCTCCAGGCTCGCGGCTCCCT	900

Db	180	TTCCAGGTTACATGAGGGGCGCCGTGTCTCCCGCAGTGTCCCAAGGCTGCCGGCTCCCT	239
Qy	901	CGACAGCTCTTGGGACACCCGGTCCCTCTCTGCCCCACCTTAGCGGCTTTGCTCCAGA	960
Db	240	CGACAGCTCTTGGGACACCCGGTCCCTCTCTGCCCCACCTTAGCGGCTTTGCTCCAGA	299
Qy	961	CTGCGCCCTCTCTAGAGGGTGGCTGGGCGCTGTTACAGTGTCTTTCCATCCACATAAT	1020
Db	300	CTGCGCCCTCTCTAGAGGGTGGCTGGGCGCTGTTACAGTGTCTTTCCATCCACATAAT	359
Qy	1021	ACAGTATCCCACTCTATCTTATTAACAATCCCCCACCACCGCCCACTCTCCACCTAGCTC	1080
Db	360	ACAGTATCCCACTCTATCTTATTAACAATCCCCCACCACCGCCCACTCTCCACCTAGCTC	419
Qy	1081	CCCAATCCCTTACCTTTGAGGCGCCCAAGTATCTCCACTCCCCCTGGCCACAGACCC	1140
Db	420	CCCAATCCCTTACCTTTGAGGCGCCCAAGTATCTCCACTCCCCCTGGCCACAGACCC	479
Qy	1141	CAGGCACTGTGTCTCAGTCTGTCTGTGGGAGAGATGGTCCAGAAAGCC	1192
Db	480	CAGGCACTGTGTCTCAGTCTGTCTGTGGGAGAGATGGTCCAGAAAGCC	531
RESULT	15		
LOCUS	BF577781	918 bp	linear
DEFINITION	602092080F1 NCL_CGAP_CO24 Mus musculus cDNA clone IMAGE:4206595 5'		
ACCESSION	BF577781		
VERSION	BF577781.1	GI:11651493	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 918)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Jeffrey E. Green, M.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LML at:		
	http://image.llnl.gov		
	Plate: LLAM967 row: P column: 20		
	High quality sequence start: 17		
	High quality sequence stop: 724.		
FEATURES	Location/Qualifiers		
source	1..918		
	/organism="Mus musculus"		
	/strain="PVB/N"		
	/db_xref="taxon:10090"		
	/clone_image="IMAGE:4206595"		
	/clone_id="NCL_CGAP_CO24"		
	/lab_host="DH10B (TI phage-resistant)"		
	/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;		
	Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.		
	Average insert size 1.6 kb. Constructed by Life		
	Technologies. Note: this is a NCL_CGAP Library."		
BASE COUNT	153 a 292 c 268 g 204 t		
ORIGIN			
Query Match	36.0%; Score 493.6; DB 12; Length 918;		
Best Local Similarity	86.9%; Pred. No. 2.2e-97;		
Matches 577; Conservative 0; Mismatches 84; Indels 3; Gaps 3;			
Qy	198	GCTGGCCGCGCCGCGCCCTCTGTCGCGCGTGCAGTTGGGAGCGCGCATCGCTGTC	257
Db	1	GCTGGCCGCGCCGCGCCCTCTGTCGCGCGTGCAGTTGGGAGCGCGCATCGCTGTC	60





	XX	10	614.6	44.8	1168	19	AAV15859	Mus musculus tumor
	XX	11	519.2	37.8	701	20	AAX23425	Mouse TRRL3 DNA.
	XX	12	245.4	17.9	282	16	AAAT22190	Human gene signatu
	XX	13	193.4	14.1	195	24	ABX29340	Colon adenocarcino
C	14	70.4	5.1	114955	20	AAK53491	Human adiponectin A1	
C	15	61	4.4	1337	20	AAI21263	Human gene expres	
C	16	60.6	4.4	1000	21	AAA02484	Human colon cancer	
C	17	60	4.4	60	24	ABN41049	Human spliced tran	
C	18	60	4.4	60	24	ABN58591	Human spliced tran	
C	19	60	4.4	60	24	ABN58593	Human spliced tran	
C	20	60	4.4	60	24	ABN58848	Human spliced tran	
C	21	60	4.4	60	24	ABN58849	Human spliced tran	
C	22	59.2	4.3	10732	21	AAA10594	Gene encoding a su	
C	23	58	4.2	1218	21	AAK02488	Human colon cancer	
C	24	54.4	4.0	600	24	ABO52496	Oligonucleotide fo	
C	25	54.4	4.0	600	24	ABO52497	Oligonucleotide fo	
C	26	54.4	4.0	1337	24	ABO68452	Listeria monocytog	
C	27	54.4	4.0	1593	21	AAA02504	Human colon cancer	
C	28	53.2	3.9	53522	24	AAD30228	Human PKD1 gene.	
C	29	53.2	3.9	53576	19	AAAT94101	Human PKD1 gene.	
C	30	53.2	3.9	53577	17	AAAT94101	Human polykystic k	
C	31	53.2	3.9	53577	19	AAAT94108	Human PDI locus b	
C	32	50	3.6	50	20	AAK56002	Human tumour necro	
C	33	50	3.6	50	21	AAA49732	Human PRO207 DNA p	
C	34	50	3.6	50	24	ABR40292	Oligonucleotide pr	
C	35	50	3.6	201	8	AAAT0195	Streptomyces prote	
C	36	50	3.6	2188	20	AAE277506	Human ovarian tumo	
C	37	49.6	3.6	1517	21	AAAT2028	cDNA encoding huma	
C	38	49.2	3.6	825	20	AAE25378	Human mFLINT #2 nu	
C	39	49.2	3.6	825	21	AAA51078	Human mature FLINT	
C	40	49.2	3.6	936	20	AAE25376	Human FLINT #2 nuc	
C	41	49.2	3.6	936	21	AAA88731	Human FMS ligand 1	
C	42	49.2	3.6	936	21	AAA51078	Human FMS coding	
C	43	49.2	3.6	936	22	AAAD07385	Human Fas ligand I	
C	44	49.2	3.6	936	22	AAF89920	Nucleotide sequenc	
C	45	49.2	3.6	936	22	AAF84738	Nucleotide sequenc	

PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.  
 XX Browning JL, Chicheportiche Y;  
 XX WPI; 1998-145619/13.  
 DR P-PSDB; AAA47525.  
 XX Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 2; pages 48-50; 69pp; English.  
 XX  
 CC The sequence is that encoding human tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. The DNA sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of prepare probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 CC  
 XX  
 SQ Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;  
 Query Match 100.0%; Score 1373; DB 19; Length 1373;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-283;  
 Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGTCAATTGTAGACATTGAAATTCGCGCGGCTCCCTCCCGATCCCTGCGG 60  
 DB 1 ATGTCAATTGTAGACATTGAAATTCGCGCGGCTCCCTCCCGATCCCTGCGG 60  
 QY 61 TCCCGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
 DB 61 TCCCGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
 QY 121 AGCCAGAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
 DB 121 AGCCAGAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
 QY 181 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGG 240  
 DB 181 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGG 240  
 QY 241 AGCGGGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 DB 241 AGCGGGGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
 QY 301 CAGGACCCGTCGGAAGTGAATCCCGAGAGAAAGAGAGAGAGAGAGAGAGAGAG 360  
 DB 301 CAGGACCCGTCGGAAGTGAATCCCGAGAGAGAAAGAGAGAGAGAGAGAGAGAG 360  
 QY 361 AACCGACTAGTTCGGCTCGCAGAGAGTGCATTAAGGCGGAAACACGCGCTCGA 420  
 DB 361 AACCGACTAGTTCGGCTCGCAGAGAGTGCATTAAGGCGGAAACACGCGCTCGA 420  
 QY 421 GCGATGCGACGCCATTTATGAATTCATCCAGACGTCGAGAGAGAGAGAGAGAG 480  
 DB 421 GCGATGCGACGCCATTTATGAATTCATCCAGACGTCGAGAGAGAGAGAGAGAG 480  
 QY 481 GTGAGAGGAGACGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 DB 481 GTGAGAGGAGACGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 541 TACACCGCAGATCGGGGAGTTTATAGTACACCGGGGCTGGCTTACTACTGTCGT 600  
 DB 541 TACACCGCAGATCGGGGAGTTTATAGTACACCGGGGCTGGCTTACTACTGTCGT 600  
 QY 601 CAGGTGACATTTGATGAGGAGGAGGCTGTCTACCTGAGCTGCTGTGTGATGCT 660  
 DB 601 CAGGTGACATTTGATGAGGAGGAGGCTGTCTACCTGAGCTGCTGTGTGATGCT 660  
 QY 661 GTGCTGGCGCTGGCGCTGGAGAGAAATTCACGACGTCGCGGCGAGTTCCCTGG 720  
 DB 661 GTGCTGGCGCTGGCGCTGGAGAGAAATTCACGACGTCGCGGCGAGTTCCCTGG 720  
 QY 721 CAGCTCGCGCTCGCAGAGGTGTGAGGCTGTGAGGCTGCGGCGAGGCTGCTGCGG 780  
 DB 721 CAGCTCGCGCTCGCAGAGGTGTGAGGCTGTGAGGCTGCGGCGAGGCTGCTGCGG 780  
 QY 781 ATCCGACCCCTCCGCGGCGCATTCAGAGCTGCGGCGCTTCCTACCTTTCGAGCT 840  
 DB 781 ATCCGACCCCTCCGCGGCGCATTCAGAGCTGCGGCGCTTCCTACCTTTCGAGCT 840  
 QY 841 TTCCAGTTTACAGAGGCGCTGGTCTCCCGACAGTCTGCGGCGGCTGCGGCTGCG 900  
 DB 841 TTCCAGTTTACAGAGGCGCTGGTCTCCCGACAGTCTGCGGCGGCTGCGGCTGCG 900  
 QY 901 CGACAGCTCTGTGGGACCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 960  
 DB 901 CGACAGCTCTGTGGGACCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 960  
 QY 961 CTTGCGGCTCTCTCTGAGAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1020  
 DB 961 CTTGCGGCTCTCTCTGAGAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1020  
 QY 1021 ACAGTATTCACCTTATCTTACAACTCCCGACGCGGCTGCGGCTGCGGCTGCGG 1080  
 DB 1021 ACAGTATTCACCTTATCTTACAACTCCCGACGCGGCTGCGGCTGCGGCTGCGG 1080  
 QY 1081 CCCAATCCCTGACCTTTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140  
 DB 1081 CCCAATCCCTGACCTTTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140  
 QY 1141 CAGGCGATTTGTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
 DB 1141 CAGGCGATTTGTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
 QY 1201 GCACATAAGAGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 DB 1201 GCACATAAGAGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
 QY 1261 TTCCCAATGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320  
 DB 1261 TTCCCAATGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320  
 QY 1321 TTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373  
 DB 1321 TTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373

RESULT 2  
 ABR34881  
 ID ABR34881 standard; cDNA; 1364 BP.  
 XX  
 AC ABR34881;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Human cDNA encoding secreted protein #19.  
 XX  
 KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;  
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;  
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;

KM coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
KW tissue regeneration; wound healing; burn; haematopoiesis;  
KM myeloid cell deficiency; lymphoid cell deficiency.  
OS Homo sapiens.  
XX MO20017288-A2.  
PN 18-OCT-2001.  
XX 29-MAR-2001; 2001WO-US10224.  
PF 06-APR-2000; 2000US-195582P.  
PR (GEM) GENETICS INST INC.  
XX Wong GG, Clark HF, Beechell K, Agostino MJ, Howes SH, Resnick RJ;  
PI Gulukota K, Graham JR;  
XX WPI; 2002-179321/23.  
DR Five hundred and ninety two polynucleotides derived from a variety of  
XX human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders  
PT  
PS Claim 1; Page 82; 372pp; English.  
XX The invention relates to 592 polynucleotides which have been derived from  
CC a variety of human tissue sources and which encode novel secreted  
CC proteins. The polynucleotides can be used as probes for the  
CC identification and isolation of full length cDNA and genomic DNA. The  
CC polynucleotides and proteins can also be used as nutritional supplements.  
CC The proteins are useful in the treatment of various immune deficiencies  
CC and disorders such as viral infections, bacterial infections, fungal  
CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
CC and conditions (e.g. asthma). They are also useful for treating  
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
CC useful for tissue regeneration, for wound healing and in the treatment of  
CC burns, incisions and ulcers. The proteins are also useful for regulating  
CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention.  
XX  
SQ Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other;  
Query Match 98.0%; Score 1345.8; DB 24; Length 1364;  
Best Local Similarity 99.9%; Pred. No. 4.3e-277;  
Matches 1347; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 25 TCCGCGCCGCGGCTCCCTCCCGATCCCTCGGCTCCGCGATGGGGGCGGTAGG 84  
DB 5 TCCGCGCCGCGGCTCCCTCCCGATCCCTCGGCTCCGCGATGGGGGCGGTAGG 64  
QY 85 CAGGACAGCGCCCGCCGCGCGCGCGCGCGCGAGAGCGGAGGGGCGCGG 144  
DB 65 CAGGACAGCGCCCGCCGCGCGCGCGCGCGCGAGAGCGGAGGGGCGCGG 124  
QY 145 GGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 204  
DB 125 GGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184  
QY 205 TGCCTGCGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 264  
DB 185 TGCCTGCGCTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 244  
QY 265 GAGCGCGCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 324  
DB 245 GAGCGCGCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 304  
QY 325 CAGACAGAAAGCAGATCTGCGCTTCTGAAACCACTAGTTGCGCTTCGAGA 384

DB 305 CAGACAGAAAGCAGATCTGCGCTTCTGAAACCACTAGTTGCGCTTCGAGA 364  
QY 385 AGTGACCTAAAGCGCGGAAAAACAGCGGCTCGAAGAGGATCGAGCCATTATGAAGT 444  
DB 365 AGTGACCTAAAGCGCGGAAAAACAGCGGCTCGAAGAGGATCGAGCCATTATGAAGT 424  
QY 445 CATCCAGACCTGGACAGAGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 504  
DB 425 CATCCAGACCTGGACAGAGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 484  
QY 505 GAGGAAGCGCAGAAATCAACAGCTCCAGCCCTGCGCTCAACCGCGCAGATGGGAGTTT 564  
DB 485 GAGGAAGCGCAGAAATCAACAGCTCCAGCCCTGCGCTCAACCGCGCAGATGGGAGTTT 544  
QY 565 ATAGTCACCGCGGCGGCGCTACTACTACTACTACTACTACTACTACTACTACT 624  
DB 545 ATAGTCACCGCGGCGGCGCTACTACTACTACTACTACTACTACTACTACTACT 604  
QY 625 GCTGCTACCTGAAGCTGAGCTGCTGATGATGATGATGATGATGATGATGATG 684  
DB 605 GCTGCTACCTGAAGCTGAGCTGCTGATGATGATGATGATGATGATGATGATG 664  
QY 685 GAATTCACGACCTGCGGCGGCGAGTTCCCTGCGGCGCGCGCTGCGCAGAGTGTCT 744  
DB 665 GAATTCACGACCTGCGGCGGCGAGTTCCCTGCGGCGCGCGCTGCGCAGAGTGTCT 724  
QY 745 GGGCTGTGGCGCTGCGGCGGCGAGGCTCCCTGCGGAGTCCGACCCCTCGGGCCAT 804  
DB 725 GGGCTGTGGCGCTGCGGCGGCGAGGCTCCCTGCGGAGTCCGACCCCTCGGGCCAT 784  
QY 805 CTCAGAGCTGCGCCCTTCTCACTACTACTACTACTACTACTACTACTACTACT 864  
DB 785 CTCAGAGCTGCGCCCTTCTCACTACTACTACTACTACTACTACTACTACTACT 844  
QY 865 GTCTCCCGCAGTGTCTCCAGGCTGCGGCTCCCTGCGAGCTCTGCGCACCCGCTC 924  
DB 845 GTCTCCCGCAGTGTCTCCAGGCTGCGGCTCCCTGCGAGCTCTGCGCACCCGCTC 904  
QY 925 CCCTGCGCCCGCAGCTTCACTACTACTACTACTACTACTACTACTACTACT 984  
DB 905 CCCTGCGCCCGCAGCTTCACTACTACTACTACTACTACTACTACTACTACT 964  
QY 985 CTGGGCGCTGTCAGCTGTTTCCATCCACATTAATACAGTATTCACCTTATCTAC 1044  
DB 965 CTGGGCGCTGTCAGCTGTTTCCATCCACATTAATACAGTATTCACCTTATCTAC 1024  
QY 1045 AACTCCCGCAGCGCTTCTCACTACTACTACTACTACTACTACTACTACTACT 1104  
DB 1025 AACTCCCGCAGCGCTTCTCACTACTACTACTACTACTACTACTACTACTACT 1084  
QY 1105 CCCAGTATCTGACCTCCCGCTGCGCCACAGACCCCGAGGGCATTTGTACTGTACTC 1164  
DB 1085 CCCAGTATCTGACCTCCCGCTGCGCCACAGACCCCGAGGGCATTTGTACTGTACTC 1144  
QY 1165 TGTGGGCAAGGATGGGTCCAGAAAGCCCGCTTCAAGGACTAAGAGGGGCTGAGCTGCG 1224  
DB 1145 TGTGGGCAAGGATGGGTCCAGAAAGCCCGCTTCAAGGACTAAGAGGGGCTGAGCTGCG 1204  
QY 1225 GGCAGGAAGCCAAAGAGACTGGGCTAGGCGCAGAGTTCCCAATGTGAGGGGCGAGAAA 1284  
DB 1205 GGCAGGAAGCCAAAGAGACTGGGCTAGGCGCAGAGGTTCCCAATGTGAGGGGCGAGAAA 1264  
QY 1285 CAAGCAAGCTCTCCCTTGAGAAATTCCTGTGATTTTAAACAGATATATTTTAT 1344  
DB 1265 CAAGCAAGCTCTCCCTTGAGAAATTCCTGTGATTTTAAACAGATATATTTTAT 1324  
QY 1345 TATTATTGTACAAATATGTGATAATG 1373  
DB 1325 TATTATTGTACAAATATGTGATAATG 1353

RESULT 3

ID	AA556000 standard: DNA: 1421 BP.
XX	AA556000;
XX	15-JUL-1999 (first entry)
DT	Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.
XX	Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
XX	NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
XX	cancer; ss.
OS	Homo sapiens.
XX	Key
XX	Location/Qualifiers
FT	92..841
FT	/*tag=
FT	/product= "Apo-3 ligand"
XX	WO919490-A1.
XX	22-APR-1999.
XX	09-OCT-1998; 98WO-US21407.
XX	17-DEC-1997; 97US-0069862.
XX	10-OCT-1997; 97US-0062037.
XX	(GETH ) GENENTECH INC.
XX	Ashkenazi AJ, Marsters SA, Pittl R;
XX	WPI: 1999-287982/24.
XX	P-PSDB: AAY09369.
XX	New human Apo-3 ligand (a tumour necrosis factor) homologue
XX	Claim 18; Fig 1; 74pp: English.
XX	The present sequence encodes a human tumour necrosis factor (TNF) and
XX	lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
XX	cytostatic activity. Apo-3 ligand can be used to induce apoptosis in
XX	mammalian cancer cells, to induce NF-kappaB-dependent transcription and
XX	to induce JNK/SAPK-dependent responses in mammalian cells.
XX	Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;
XX	Query Match 96.5%; Score 1325; DB 20; Length 1421;
XX	Best Local Similarity 98.5%; Pred. No. 1.2e-272;
XX	Matches 1337; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY	17 TTGAAATTTCGGCCGGCGCTCCCTCCCGATCCCTCGGGTCCGGGATGGGGG 76
DB	3 TTAGAGTTCCTCGACTCGACCCAGCCGTCGCGATCCCTCGGGTCCGGGATGGGGG 62
QY	77 CGGTGAGCAGACAGCCCGCCCGCCATGAGCGCGCGTGGAGACAGAGCGGAGGG 136
DB	63 CGGTGAGCAGACAGCGCCCGCCCGCCATGAGCGCGCGTGGAGACAGAGCGGAGGG 122
QY	137 GCGCGCGGGGAGCCGGGACCGCCCTGCTGCTCCGCTCGCGCTGGGCTGGGCTGG 196
DB	123 GCGCGCGGGGAGCCGGGACCGCCCTGCTGCTCCGCTCGCGCTGGGCTGGGCTGG 182
QY	197 CGCTGGCTTCGCTCGGCTTCCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
DB	183 CGCTGGCTTCGCTCGGCTTCCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
QY	257 CGCGCCAGAGAGCTGGCCAGAGAGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 316
DB	243 CGCGCCAGAGAGCTGGCCAGAGAGAGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAG 302
QY	317 TGAATTCGCGAGAGAGAAAGCCAGATCTGCGCTTCTCTGAACGACTAGTTGGGC 376

Db	303	TGATGCCAGACAGAAAGAGACAGATCTCGCGCTTCTCTGAACCGACTAGTTGGGC	362
Qy	377	CTCCAGAGTGCACCTAAAGGCCGGAACACAGGGCTTGAAGAGCATCCAGCCATT	436
Db	363	CTCCAGAAATGACACCTAAAGGCCGGAACACAGGGCTTGAAGAGCATCCAGCCATT	422
Qy	437	ATGAAGTTCATCCAGCACTGGACAGAGAGAGCCAGGACAGTGTGGACGGAGACGA	496
Db	423	ATGAAGTTCATCCAGCACTGGACAGAGAGAGCCAGGACAGTGTGGACGGAGACGA	482
Qy	497	GTCGCTGGAGGAAGCCAGATCAACAGCTCCAGCCCTTCGCTACACACCGCAGATCG	556
Db	483	GTCGCTGGAGGAAGCCAGATCAACAGCTCCAGCCCTTCGCTACACACCGCAGATCG	542
Qy	557	GGGAGTTATAGTACACCCGGGGCTGGGCTCTACTACCTGTACTGTGTAAGGACATTGATG	616
Db	543	GGGAGTTATAGTACACCCGGGGCTGGGCTCTACTACCTGTACTGTGTAAGGACATTGATG	602
Qy	617	AGGGGAAGGCTGTCTACCTGAAGCTGGACTGTGCTGTGATGGTGTGCGCCGCGCT	676
Db	603	AGGGGAAGGCTGTCTACCTGAAGCTGGACTGTGCTGTGATGGTGTGCGCCGCGCT	662
Qy	677	GCCCTGAGAGAAATTCTACGCCACTGCGGCCAGTTCCTCGGGGCCCAAGCTCGGCTCTGCC	736
Db	663	GCCCTGAGAGAAATTCTACGCCACTGCGGCCAGTTCCTCGGGGCCCAAGCTCGGCTCTGCC	722
Qy	737	AGGTGTCTGGGCTTTTGGCCCTGCGGCCAAGGGTCCTCCCTGGGAGATCGCACACCTCCCT	796
Db	723	AGGTGTCTGGGCTTTTGGCCCTGCGGCCAAGGGTCCTCCCTGGGAGATCGCACACCTCCCT	782
Qy	797	GGGCCCATCTCAGGCTGGCCCTTCTCACTACTTGGGACTTTCCAGAGTTACTATAG	856
Db	783	GGGCCCATCTCAGGCTGGCCCTTCTCACTACTTGGGACTTTCCAGAGTTACTATAG	842
Qy	857	GGGCCCTGGTCTCCCAAGTCTGCCAGGCTCGCGGCTCCCTCGACAGACTCTTGGGC	916
Db	843	GGGCCCTGGTCTCCCAAGTCTGCCAGGCTCGCGGCTCCCTCGACAGACTCTTGGGC	902
Qy	917	ACCCGGTCCCTCTGCCCAACCTCAGCGGCTTTTGTCTCCAGAGCTCGCCCTCCCTCTA	976
Db	903	ACCCGGTCCCTCTGCCCAACCTCAGCGGCTTTTGTCTCCAGAGCTCGCCCTCCCTCTA	962
Qy	977	GAGGCTGCTGGGCTGTTCACGTGTTTTCATCCACATTAATACAGTATCCCACTCT	1036
Db	963	GAGGCTGCTGGGCTGTTCACGTGTTTTCATCCACATTAATACAGTATCCCACTCT	1022
Qy	1037	TATCTTAACAATCCCCACCGCCACTCTTCACACTAGCTAGCTCCCAATCCCTGACCT	1096
Db	1023	TATCTTAACAATCCCCACCGCCACTCTTCACACTAGCTAGCTCCCAATCCCTGACCT	1082
Qy	1097	TTTGAAGCCCCCAGTATCTGACCTCCGCCCTGGGCAACAGCCCCAGGGCAATTTGTGTCA	1156
Db	1083	TTTGAAGCCCCCAGTATCTGACCTCCGCCCTGGGCAACAGCCCCAGGGCAATTTGTGTCA	1142
Qy	1157	CTGTACTCTGTGGGCAAGATGGGTCCAGAAAGCCCACTTAGCGCACTAAGAGGGGCTG	1216
Db	1143	CTGTACTCTGTGGGCAAGATGGGTCCAGAAAGCCCACTTAGCGCACTAAGAGGGGCTG	1202
Qy	1217	GACCTGGCGCAGAGAACCAAAAGACATGGGCTAGGCCAGGAGTTCCCAATGTGAAGG	1276
Db	1203	GACCTGGCGCAGAGAACCAAAAGACATGGGCTAGGCCAGGAGTTCCCAATGTGAAGG	1262
Qy	1277	GCGAGAAACAAGACAACCTCTCCCTTTGAAGTAATTCCTGTGGATTTTAAACACATATT	1336
Db	1263	GCGAGAAACAAGACAACCTCTCCCTTTGAAGTAATTCCTGTGGATTTTAAACACATATT	1322
Qy	1337	ATTTTATTTATTTATGTGACAAAATGTTGATTAATGG	1373
Db	1323	ATTTTATTTATTTATTTGACAAAATGTTGATTAATGG	1359

RESULT 4

AAA49717  
 ID AAA49717 standard; cDNA; 1353 BP.  
 AC AAA49717;  
 XX  
 XX 25-SEP-2000 (first entry)  
 XX  
 XX Human PRO207 cDNA clone DNA30879-1152.  
 XX  
 XX PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KM breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KM uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KM central nervous system cancer; melanoma; leukaemia; neoplasm; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 58..807  
 FT sig\_peptide /\*tag= a  
 FT 58..177  
 FT /\*tag= b  
 FT mat\_peptide 178..804  
 FT /\*tag= c  
 XX  
 XX MO200037638-A2.  
 PD 29-JUN-2000.  
 XX  
 PF 02-DEC-1999; 99MO-US28565.  
 XX  
 PR 22-DEC-1998; 98US-0113296.  
 PR 08-MAR-1999; 99MO-US05028.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99MO-US21090.  
 PR 15-SEP-1999; 99MO-US21547.  
 XX  
 PA (GETH ) GENENTECH INC.  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pitti RM, Wood WI;  
 XX  
 DR WPI: 2000-442668/38.  
 DR P-PSDB; AAY95338.  
 XX  
 XX Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or  
 PT PRO866  
 PS  
 PS Claim 20; Fig 3; 172pp; English.  
 XX  
 XX The present sequence is that of cDNA clone DNA30879-1152  
 CC (ATCC 20335) encoding human PRO207 (see AAY95338), which shows  
 CC homology to several members of the tumour necrosis factor family,  
 CC especially human lymphotoxin (23.48). The cDNA was identified in a  
 CC foetal kidney cDNA library following identification of an expressed  
 CC sequence tag with homology to human Apo-2 ligand. A claimed method  
 CC for inhibiting the growth of a tumour cell comprises exposing the  
 CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,  
 CC PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see  
 CC AAY95337-49), their agonists or chimeric polypeptides incorporating  
 CC them. The tumour is especially a cancer selected from breast,  
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and  
 CC central nervous system cancer, melanoma and leukaemia. Nucleic  
 CC acids encoding PRO179 etc. are used in the recombinant production  
 CC of the antitumour polypeptides.  
 XX  
 XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

Query Match 96.2%; Score 1320.2; DB 21; Length 1353;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-271;  
 Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 49 CGATCCCTCGGGTCCGGGATGGGGGGGCTGAGCAGCAGCAGCCCGCCCATG 108  
 DB 1 CGATCCCTCGGGTCCGGGATGGGGGGGCTGAGCAGCAGCAGCCCGCCCATG 60  
 QY 109 GCGGCGCGTGGAGCCAGAGCCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 168  
 DB 61 GCGGCGCGTGGAGCCAGAGCCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 120  
 QY 169 GTCCCGCTCGGCGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG 228  
 DB 121 GTCCCGCTCGGCGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGG 180  
 QY 229 GTCAATTTGGGGAGCCGGGATCGCTGTCCGCCAGAGCCCTGCCAGAGAGCTGGT 288  
 DB 181 GTCAATTTGGGGAGCCGGGATCGCTGTCCGCCAGAGCCCTGCCAGAGAGCTGGT 240  
 QY 289 GCAGAGAGAGACAGAGACCCGTCGGAATCGAATCCCAAGAGAGAGAGAGATCT 348  
 DB 241 GCAGAGAGAGACAGAGACCCGTCGGAATCGAATCCCAAGAGAGAGAGATCT 300  
 QY 349 GCGCCTTTCCTGAGCCGACTAGTTCGGGCTCGCAGAAAGTGCACCTAAAGCGGAAAA 408  
 DB 301 GCGCCTTTCCTGAGCCGACTAGTTCGGGCTCGCAGAAAGTGCACCTAAAGCGGAAAA 360  
 QY 409 CGGGCTCGAAGAGAGATCGAGCCCATTTATGACCTATCCACAGCCTGACAGAGAG 468  
 DB 361 CGGGCTCGAAGAGAGATCGAGCCCATTTATGACCTATCCACAGCCTGACAGAGAG 420  
 QY 469 GCGCAGCAGAGTGTGAGAGCGGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGTCC 528  
 DB 421 GCGCAGCAGAGTGTGAGAGCGGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGTCC 480  
 QY 529 AGCCCTTTCGCGCTACACCGCCAGATGGGGAGTTTATGACCCCGGCTGGCTCTAC 588  
 DB 481 AGCCCTTTCGCGCTACACCGCCAGATGGGGAGTTTATGACCCCGGCTGGCTCTAC 540  
 QY 589 TACCTGTACTGTCAAGTGTCACTTTGATGAGGGAGAGCTGTCTACCTGAAAGCTGG 648  
 DB 541 TACCTGTACTGTCAAGTGTCACTTTGATGAGGGAGAGCTGTCTACCTGAAAGCTGG 600  
 QY 649 CTGGTGATGGTGTGCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 708  
 DB 601 CTGGTGATGGTGTGCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 660  
 QY 709 TCCCTCGGGCCCGAGCTCCGCTCTGCGAGGTCTGTGGGCTGTGGGCTGTGGGCGAGG 768  
 DB 661 TCCCTCGGGCCCGAGCTCCGCTCTGCGAGGTCTGTGGGCTGTGGGCTGTGGGCGAGG 720  
 QY 769 TCTTCCTCGGGATCGGACCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTCTACC 828  
 DB 721 TCTTCCTCGGGATCGGACCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTCTACC 780  
 QY 829 TACTTCGAGCTTCCAGGTTCACATGAGGGGGCGCTGGGCTGCCAGATGTGCCAGGCT 888  
 DB 781 TACTTCGAGCTTCCAGGTTCACATGAGGGGGCGCTGGGCTGCCAGATGTGCCAGGCT 840  
 QY 889 GCGGCGTCCCTCGAGACGCTCTGAGGACCGGGTCCCTCTGCGCCACCTCAGCGCGCT 948  
 DB 841 GCGGCGTCCCTCGAGACGCTCTGAGGACCGGGTCCCTCTGCGCCACCTCAGCGCGCT 900  
 QY 949 CTTTGTCTCAGACTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1008  
 DB 901 CTTTGTCTCAGACTGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960  
 QY 1009 TCCGACATTAATACAGTATTCACCTCTATCTTACAACTCCGACGCGGACCTCTTCCA 1068  
 DB 961 TCCGACATTAATACAGTATTCACCTCTATCTTACAACTCCGACGCGGACCTCTTCCA 1020  
 QY 1069 CCTCAGTACGTCCCAATCCCTGACCTTTGAGGCGCCCGCAGTGTCTGACTCCCGCTG 1128



```
|||||
Db 721 TCTCTCCGCGGATCGGACCGCTCCCTGGGCGCCATCTCAAGCGTCCCGCTTCTCTCAC 780
Qy 829 TACTTGGACTCTTCCAGGTTCTACTAGAGGGCCCTGTCTCCCAAGTGTGCCAGGCT 888
Db 781 TACTTGGAGCTTCTCAGGTTCTACTAGAGGGCCCTGTCTCCCGAGTGTGCCAGGCT 840
Qy 889 GCGGCGTCCCTCGAGACTCTCTGCGGACCGCGTCCCTGCGCCCGCCAGCCGCT 948
Db 841 GCGGCGTCCCTCGAGACTCTCTGCGGACCGCGTCCCTGCGCCCGCCAGCCGCT 900
Qy 949 CTCTGCTCCAGACCTGCGCCCTCTCTAGAGGCTGCTGGGCTTTGACGTGTTTCCA 1008
Db 901 CTCTGCTCCAGACCTGCGCCCTCTCTAGAGGCTGCTGGGCTTTGACGTGTTTCCA 960
Qy 1009 TCCCACTAATAATACAGTATTTCCCACTTTATCTTACAACTCCCGCCAGCTCTCCA 1068
Db 961 TCCCACTAATAATACAGTATTTCCCACTTTATCTTACAACTCCCGCCAGCTCTCCA 1020
Qy 1069 CCTCACTAGCTCCCAATCCCTGAGCCCTTTGAGGCGCCCGAGATCTGACTCCCGCTG 1128
Db 1021 CCTCACTAGCTCCCAATCCCTGAGCCCTTTGAGGCGCCCGAGATCTGACTCCCGCTG 1080
Qy 1129 GCCACAGACCCCGAGGATGTGTCTACTCTGTGCGCAAGATGGGTCCAGAAG 1188
Db 1081 GCCACAGACCCCGAGGATGTGTCTACTCTGTGCGCAAGATGGGTCCAGAAG 1140
Qy 1189 ACCCACTTCAGGACTAAGAGGGGCTGAGCTGGCGGAGAGCCAAAGACTGGGC 1248
Db 1141 ACCCACTTCAGGACTAAGAGGGGCTGAGCTGGCGGAGAGCCAAAGACTGGGC 1200
Qy 1249 CTAGCGCAGAGTTCCCAATGTAGGGGCGAGAAACAAGACGCTCCCTTGAGAA 1308
Db 1201 CTAGCGCAGAGTTCCCAATGTAGGGGCGAGAAACAAGACGCTCCCTTGAGAA 1260
Qy 1309 TTCCCTGTGATTTTAAACAGATATATTTTATTTATTTGACAAAATGTTGAA 1368
Db 1261 TTCCCTGTGATTTTAAACAGATATATTTTATTTATTTGACAAAATGTTGAA 1320
Qy 1369 AATGG 1373
Db 1321 AATGG 1325

RESULT 6
AAV47613
ID AAV47613 standard; cDNA: 1236 BP.
AC AAV47613;
XX
DT 27-OCT-1998 (first entry)
XX
DE TNF related endothelium proliferative agent gene.
XX
KW ss: TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..750
FT /tag= a
FT /product= "TREPA"
XX
PM W09835061-A2.
XX
PD 13-AUG-1998.
XX
PF 12-FEB-1998; 98MO-US02859.
XX
PR 10-FEB-1998; 98US-0021706.
XX
PR 12-FEB-1997; 97US-0798692.
```

```
PA (ABBO ) ABBOTT LAB.
XX
PI WILEY SR:
XX
DR WPI: 1998-447255/38.
DR P-PSDB: AAM29745.
XX
PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
PT treatment of autoimmune disease, tumours and inflammation
XX
PS Claim 11; Page 123-4; 142pp; English.
XX
CC The TNF-related endothelium proliferative agent (TREPA), or its
CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
CC promote wound healing or tissue grafting, by promoting vascularisation,
CC also to induce apoptosis for treating cancer and eliminating autoreactive
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
CC TREPA peptides can also be used to target cytotoxic agents or for
CC affinity isolation of the corresponding receptor, the nucleic acid for
CC which can be used to transform tumour cells to render them more
CC responsive to TREPA and to screen for TREPA mimics.
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
CC vascularisation), inflammation or a wide range of autoimmune conditions,
CC conditions involving abnormal stimulation of epithelial cells (e.g.
CC atherosclerosis), for birth control (inhibiting ovulation and placental
CC formation) or other angiogenic conditions (e.g. ulcers).
XX
SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Query Match 89.3%; Score 1226.4; DB 19; Length 1236;
Best Local Similarity 99.5%; Pred. No. 1.1e-251;
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 106 ATGCGCGCCGTCGAGACCGAGGCGGCGCGGCGGAGACCGCGGCGCTG 165
Db 1 ATGCGCGCCGTCGAGACCGAGGCGGCGGCGGCGGAGACCGCGGCGCTG 60
Qy 166 CTGGTCCCGCTGCGGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 225
Db 61 CTGGTCCCGCTGCGGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 120
Qy 226 GTGTGTCAGTTTGGGAGCGCGGCGATCGCTGCGGCCAGAGACCTGCCAGAGAGCTG 285
Db 121 GTGTGTCAGTTTGGGAGCGCGGCGATCGCTGCGGCCAGAGACCTGCCAGAGAGCTG 180
Qy 286 GTGCGAGAGAGACCGAGACCGCTGCGAATGATCCCGACAGAGAGAAAGCCAGAT 345
Db 181 GTGCGAGAGAGACCGAGACCGCTGCGAATGATCCCGACAGAGAGAAAGCCAGAT 240
Qy 346 CCGTGGCCCTTCTGTAACGATAGTGGCCCTCGAGAAAGCAACCTAAAGCGCGGAA 405
Db 241 CCGTGGCCCTTCTGTAACGATAGTGGCCCTCGAGAAAGCAACCTAAAGCGCGGAA 300
Qy 406 ACACGGGCTCGAAGACGATCGCAGCCCATTTAAGATTTCATCCAGACTGGACAGAC 465
Db 301 ACACGGGCTCGAAGACGATCGCAGCCCATTTAAGATTTCATCCAGACTGGACAGAC 360
Qy 466 GGAGCGCAGGACAGGTGTGAGCGAGACAGTGTGCTGGGAGAGAACCAATCAACAGC 525
Db 361 GGAGCGCAGGACAGGTGTGAGCGAGACAGTGTGCTGGGAGAGAACCAATCAACAGC 420
Qy 526 TCGAGCCCTGCGGTACAAACCGCAAGTCGGGAGTTATATGTCACCGGGCTGGGCTC 585
Db 421 TCGAGCCCTGCGGTACAAACCGCAAGTCGGGAGTTATATGTCACCGGGCTGGGCTC 480
Qy 586 TACTACTGTACTGTAGCTGACGCTTTGATGAGGGGAAAGCTGTACTGTGAAGCTGAG 645
Db 481 TACTACTGTACTGTAGCTGACGCTTTGATGAGGGGAAAGCTGTACTGTGAAGCTGAG 540
Qy 646 TTGCTGTGTGATGTGTGTGCTGCGCTGCTGCTGCTGAGAAATTTCTACGACCTGCGGC 705
Db 541 TTGCTGTGTGATGTGTGTGCTGCGCTGCTGCTGCTGAGAAATTTCTACGACCTGCGGC 600
```

	Y	706	ACTTCCCTGGGGCCAGCTCCGCTCAGCTCAGAGTCTCGGGCTGTGGCCCTGGGGCCA	765
OY	706	ACTTCCCTGGGGCCAGCTCCGCTCAGCTCAGAGTCTCGGGCTGTGGCCCTGGGGCCA	765	
Db	601	AGTTCCCTGGGGCCAGCTCCGCTCAGCTCAGAGTCTCGGGCTGTGGCCCTGGGGCCA	660	
OY	766	GGGTCCCTCCCTCGAGTCCGACACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTC	825	
Db	661	GGGTCCCTCCCTCGAGTCCGACACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTC	720	
OY	826	ACCTACTTGGAGCTCTTCCAGGTTACTACTAGGGGCCCTGTCTCCCAAGTGTCCAG	885	
Db	721	ACCTACTTGGAGCTCTTCCAGGTTACTACTAGGGGCCCTGTCTCCCAAGTGTCCAG	780	
OY	886	GCTGCCGGCTCCCTCGAGACGCTCTCTGGGACACCCGGTCCCTCTGCCACCTCAGCC	945	
Db	781	GCTGCCGGCTCCCTCGAGACGCTCTCTGGGACACCCGGTCCCTCTGCCACCTCAGCC	840	
OY	946	GCTCTTTGGTCCAGACCTGCGCCCTCCCTCTAGAGGCTGCCGTGGCCCTGTTCAAGTGT	1005	
Db	841	GCTCTTTGGTCCAGACCTGCGCCCTCCCTCTAGAGGCTGCCGTGGCCCTGTTCAAGTGT	900	
OY	1006	CCATCCCACTAAATACAGTATTCCCACTCTTATCTTACAACCTCCCCACCGCCACTCT	1065	
Db	901	CCATCCCACTAAATACAGTATTCCCACTCTTATCTTACAACACCCACCGCCACTCT	960	
OY	1066	CCACCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCCCCAGTATCTGCATCCCCC	1125	
Db	961	CCACCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCCCCAGTATCTGCATCCCCC	1020	
OY	1126	CTGGCCACAGACCCCCAGGGCAATTGTGTCACTGACTCTGTGGCAGAGATGGTCCAG	1185	
Db	1021	CTGGCCACAGACCCCCAGGGCAATTGTGTCACTGACTCTGTGGCAGAGATGGTCCAG	1080	
OY	1186	AGACCCCACTCAGGCACTAGAGGGGCTGGACCTGGGGGCGAGAGCCAAAGAGACTG	1245	
Db	1081	AGACCCCACTCAGGCACTAGAGGGGCTGGACCTGGGGGCGAGAGCCAAAGAGACTG	1140	
OY	1246	GGCTAGGCGCAGAGTTCCTCAATGTGAGAGGGCGAGAAACAAGACCTCCTCTTGA	1305	
Db	1141	GGCTAGGCGCAGAGTTCCTCAATGTGAGAGGGCGAGAAACAAGACCTCCTCTTGA	1200	
OY	1306	GAATTCCCTGTGGATTTTAAACAGATATATTTT	1341	
Db	1201	GAATTCCCTGTGGATTTTAAACAGATATATTTT	1236	
RESULT 7				
ID	AAD04350			
AD	AAD04350	standard; cDNA; 1236 BP.		
XX	AAD04350;			
XX				
DT	04-JUL-2001	(first entry)		
XX				
DE	Human TREPA (TNF related endothelium proliferative agent) cDNA.			
XX				
KM	Human; tumour necrosis factor; TNF; angiogenesis; wound healing;			
KM	TREPA; TNF related endothelium proliferative agent; tumour; metastasis;			
KM	grafting; vulnerability; ss.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FH	CDS	1..750		
FT		/*tag= a		
FT		/product= "Human TREPA (TNF related endothelium		
XX		proliferative agent)"		
XX	US6207642-B1.			
XX				
PD	27-MAR-2001.			
XX				
PF	26-JUN-1998; 98US-0105343.			

XX 12-FEB-1997; 97US-0798692.  
PR 10-FEB-1998; 98US-0021706.  
XX  
XX (ABBO ) ABBOTT LAB.  
XX  
XX WILEY SR;  
XX  
XX WPI; 2001-280760/29.  
DR  
DR P-PSDB; AAE00891.  
XX  
XX  
XX Inducing angiogenesis in mammal at desired sites for promoting wound  
PT healing, by administering soluble fragment of extracellular domain of  
PT tumor necrosis factor related endothelium proliferative agent protein  
PT -  
XX  
XX  
XX Example 2; Column 73-74; 53pp; English.  
XX  
XX The present invention relates to extracellular signal molecules,  
CC particularly members of tumour necrosis factor (TNF) family molecules  
CC designated as TREPA (TNF related endothelium proliferative agent) .  
CC Soluble biologically active TREPA are used to treat TREPA-associated  
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
CC in human for promoting wound healing and for vascularising grafted  
CC tissue for successful grafting and to promote tissue grafts.  
CC The present sequence is a cDNA clone ID #69050 encoding human TREPA.  
XX  
XX  
XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

[illegible]



```
QY 706 AGTTCCTCGGGGCCCCAGCTCCGCTCTGCGAGGTGTGTGGCTGTGTGGCCCTGGGGGCA 765
    |||||||
Db 601 AGTTCCTCGGGGCCCCAGCTCCGCTCTGCGAGGTGTGTGGCTGTGTGGCCCTGGGGGCA 660
QY 766 GGGTCTCTCCCTGGGAGTCCGACCCCTCCCTGGGGCCATCTCAAGGTGCCCCCTTCTCTC 825
    |||||||
Db 661 GGGTCTCTCCCTGGGAGTCCGACCCCTCCCTGGGGCCATCTCAAGGTGCCCCCTTCTCTC 720
QY 826 ACCACTTCGGAGCTCTCCAGGTTCACATGAGGGGGCTGGGTCTCCCAACATCTGGCCAG 885
    |||||||
Db 721 ACCACTTCGGAGCTCTCCAGGTTCACATGAGGGGGCTGGGTCTCCCAACATCTGGCCAG 780
QY 886 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCCGGTCTCCCTGCCCCACCTCAGCC 945
    |||||||
Db 781 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCCGGTCTCCCTGCCCCACCTCAGCC 840
QY 946 GCTCTTTGCTCCAGACCTGCCCCCTCTCTAGAGGCTCCCTGGGCTGTTCACGTGTCTT 1005
    |||||||
Db 841 GCTCTTTGCTCCAGACCTGCCCCCTCTCTAGAGGCTCCCTGGGCTGTTCACGTGTCTT 900
QY 1006 CCATCCCAATAAATACAGTATTCCTCACTTATCTTCAACTCCGCCACGCCCACTCT 1065
    |||||||
Db 901 CCATCCCAATAAATACAGTATTCCTCACTTATCTTCAACAACCCACGCCCACTCT 960
QY 1066 CCACCTCACTAGCTCCCAATCCCTGACCCCTTTGAGGGCCCCAGTATCTGACTCCGCC 1125
    |||||||
Db 961 CCACCTCACTAGCTCCCAATCCCTGACCCCTTTGAGGGCCCCAGTATCTGACTCCGCC 1020
QY 1126 CTGGCCACAGACCCCGAGGAGTGTGTCTACTGTCTGTGGGCAAGATGGGTCCAG 1185
    |||||||
Db 1021 CTGGCCACAGACCCCGAGGAGTGTGTCTACTGTCTGTGGGCAAGATGGGTCCAG 1080
QY 1186 AAGACCCCACTTCGAGGCACTAAGAGGGGCTGGACCTGGGCGAGGAACCAAGAGACTG 1245
    |||||||
Db 1081 AAGACCCCACTTCGAGGCACTAAGAGGGGCTGGACCTGGGCGAGGAACCAAGAGACTG 1140
QY 1246 GGCCTAGGCGCAGAGTTCCTCAATGTGAGGGGCGAGAAACAAGACAGCTCTCCCTTGA 1305
    |||||||
Db 1141 GGCCTAGGCGCAGAGTTCCTCAATGTGAGGGGCGAGAAACAAGACAGCTCTCCCTTGA 1200
QY 1306 GAATTCCTCTGTGATTTTAAACGATATATTTT 1341
    |||||||
Db 1201 GAATTCCTCTGTGATTTTAAACGATATATTTT 1236

RESULT 8
AAx23424
ID AAx23424 standard; DNA; 1030 BP.
XX
AC AAx23424;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human TNRL3 DNA.
XX
KW Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;
    developmental abnormality; gestational abnormality; prostate cancer;
    AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
    cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
    apoptosis; human; ss.
XX
KM
XX
OS Homo sapiens.
XX
FH Key 1..627 Location/Qualifiers
FT CDS /*tag= a
FT /product= "TNRL3"
XX
PN MO9911791-A2.
XX
PD 11-MAR-1999.
XX
PF 04-SEP-1998; 98WO-US18393.
```

```
XX
PR 05-SEP-1997; 97US-0924634.
PA (UNIV ) UNIV WASHINGTON.
XX
PI Chaudhary PM;
XX
DR WPI: 1999-205191/17.
DR P-PSDB: AAM93590.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
    useful for diagnosis and treatment of prostate cancer and
    developmental or gestational abnormalities
XX
PS Example VII; Fig 13A; 156pp; English.
XX
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
    receptor polypeptides: AP04, AP06, AP08 and AP09 or their active
    fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
    their active fragments. AP04 is useful for diagnosing prostate cancer
    by determining levels of AP04 in an individual. Prostate cancer can also
    be treated using AP04 selective binding agents linked to a therapeutic
    moiety. AP04 polypeptides are also useful for identifying selective
    binding agents, useful in diagnosis/treatment of disease by binding of
    agents to the polypeptide/active fragment which is extracellular, or
    expressed on the cell surface. The binding is preferably performed in
    vivo. AP04 polypeptides/active fragments are also useful for screening
    for agonists and antagonists by binding and observing the change in AP04
    activity. Effective pharmacological agents useful in diagnosis or
    treatment of disease are also identified using AP04 polypeptides/active
    fragments and AP04 signal transducer molecules that specifically interact
    with a cytoplasmic domain of AP04 and detecting a change in level of AP04
    activity. The method is performed in vivo or in vitro. APO polypeptides
    are all useful as immunogens for preparing antibodies. AP04 is also
    useful for diagnosis/treatment of developmental or gestational
    abnormalities. AP08 was transfected to human breast carcinoma cell line
    MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;
    Query Match 60.7%; Score 833.4; DB 20; Length 1030;
    Best Local Similarity 99.9%; Pred. No. 4.8e-168;
    Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 229 GTCAAGTTGGGAGCGGGGCAATCCGCTGTCGGCCAGAGAGCCGCCAGAGAGCTGTG 288
    |||||||
Db 1 GTCAAGTTGGGAGCGGGGCAATCCGCTGTCGGCCAGAGAGCTGTG 60
QY 289 GCAGAGAGAGACAGACCCCGTCGAGACTGAATCCCAAGACAGAAAGACAGAGATCCT 348
    |||||||
Db 61 GCAGAGAGAGACAGACCCCGTCGAGACTGAATCCCAAGACAGAAAGACAGAGATCCT 120
QY 349 GCGCCTTTCCTGAACCGACTAGTTGGCCTGCGACAAGTGCACCTTAAAGCCGGAAAAACA 408
    |||||||
Db 121 GCGCCTTTCCTGAACCGACTAGTTGGCCTGCGACAAGTGCACCTTAAAGCCGGAAAAACA 180
QY 409 CGGCGTGAAGAGGATGCGAGGCCCATATGAAAGTTCATCCACGACCTCGAGAGAGAGGA 468
    |||||||
Db 181 CGGCGTGAAGAGGATGCGAGGCCCATATGAAAGTTCATCCACGAGCTGGAGAGAGAGGA 240
QY 469 GCGCAGGAGAGTGTGAGCGGAGACATGAGTGGCTGGAGAGAAAGCCAGATCAACAGCTCC 528
    |||||||
Db 241 GCGCAGGAGAGTGTGAGCGGAGACATGAGTGGCTGGAGAGAAAGCCAGATCAACAGCTCC 300
QY 529 AGCCCTGTGCGCTACACCCCGAGATCGGGAGATTATAGTACCCCGGGCTGAGCTTAC 588
    |||||||
Db 301 AGCCCTGTGCGCTACACCCCGAGATCGGGAGATTATAGTACCCCGGGCTGAGCTTAC 360
QY 589 TACCTGTACTGTGACAGTGCACCTTTGATAGAGGAGAGGCTGTACTTAAGCTGACATG 648
    |||||||
Db 361 TACCTGTACTGTGACAGTGCACCTTTGATAGAGGAGAGGCTGTACTTAAGCTGACATG 420
QY 649 CTGGTGTAGTGTGTGCTGGCCCTGCGCTGCGTGGAGGAATTCAGCCACTGCGGCAGT 708
```

```

DB 421 CTGCTGATGATGCTGCTGCTGCTGCTGAGAGAAATTCACAGCCACTGCGGCGCAGT 480
OY 709 TCCCTGCGGCGCCAGCTGCGGCTGCGGCGGCTGCTGCGGCTGCTGCGGCGGCGG 768
DB 481 TCCCTGCGGCGCCAGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCGGCGG 540
OY 769 TCCCTGCGGCGGATCCGACACCTCCCTGCGGCGGCTGCTGCGGCGGCTGCTGCGG 828
DB 541 TCCCTGCGGCGGATCCGACACCTCCCTGCGGCGGCTGCTGCGGCGGCTGCTGCGG 600
OY 829 TACTGCGACTCTTCCAGCTTCACTGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCT 888
DB 601 TACTGCGACTCTTCCAGCTTCACTGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCT 660
OY 889 GCGGCGTCCCTGCGACAGCTGCTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
DB 661 GCGGCGTCCCTGCGACAGCTGCTGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
OY 949 CTCTGCTCCAGACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008
DB 721 CTCTGCTCCAGACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
OY 1009 TCCCATATATATACATATATCCACCTTATCTTACAACTCCCGCCAGCGGCGC 1063
DB 781 TCCCATATATATACATATATCCACCTTATCTTACAACTCCCGCCAGCGGCGC 835

RESULT 9
AAS03964
ID AAS03964 standard; DNA: 898 BP.
XX
AC AAS03964;
XX
DT 26-SEP-2001 (first entry)
XX
DE Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
XX
KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
KW rubrosis; uveitis; macular degeneration; arthritis; rheumatism; ds;
KW corneal graft neovascularisation; psoriasis; benign tumour; haemophilic joint;
KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
KW preneoplastic condition; myocardial angiogenesis; wound granulation;
KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;
KW peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
KW fusion protein.
XX
OS Homo sapiens.
XX
XX Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 52..873
FT /tag= "Fusion protein comprising a growth hormone
FT /product= leader, a leucine zipper multimerisation
FT domain, and human TWEAK extracellular
FT domain"
XX
PN WO200145730-A2.
XX
PD 28-JUN-2001.
XX
PF 19-DEC-2000; 2000WO-US34755.
XX
PR 20-DEC-1999; 99US-0172878.
XX
PR 10-MAY-2000; 2000US-0203347.
XX
PA (IMMUNEX) IMMUNEX CORP.
XX
PI Wiley SR;

```

```

XX WPI: 2001-417975/44.
DR P-PSDB: AAU03499.
XX
PT Modulating angiogenesis in a mammal for treating diseases mediated by
PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
PT peripheral tissue, by administering antagonist or agonist of TWEAK
PT receptor
XX
PS Example 1; Page 39-40; 46pp: English.
XX
CC The sequence represents a DNA from the expression vector
CC pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth
CC hormone leader, a leucine zipper multimerisation domain, and the
CC extracellular domain of human TWEAK. The fusion protein was used in
CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones
CC from a COS cell human cDNA library. The TWEAK protein is a
CC member of the tumour necrosis factor (TNF) family and induces
CC angiogenesis. TWEAKR may therefore be used to screen for and develop
CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
CC used in the treatment and diagnosis of human disease. The disorders
CC mediated by angiogenesis include ocular disorders characterised by ocular
CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,
CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
CC rubrosis, uveitis, macular degeneration and corneal graft
CC neovascularisation, and inflammatory diseases such as arthritis,
CC rheumatism and psoriasis. Other treatable diseases include malignant and
CC metastatic conditions such as sarcomas and carcinomas, benign tumours and
CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
CC scleroderma, vascular adhesions, atherosclerotic plaque
CC neovascularisation, telangiectasia, wound granulation, coronary
CC atherosclerosis, peripheral atherosclerosis and ischaemia.
CC
SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;
XX
XX
Query Match 45.8%; Score 629.2; DB 22; Length 898;
Best Local Similarity 99.5%; Pred. No. 1,4e-124;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 232 AGTTTGGGAGCGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 291
DB 250 AGTTTGGGAGCGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
OY 292 GAGGAGACCGAGACCGCGTGGAGTGAATCCCGACAGAAAGAACCGAGATCCTGGC 351
DB 310 GAGGAGACCGAGACCGCGTGGAGTGAATCCCGACAGAAAGAACCGAGATCCTGGC 369
OY 352 CCTTCTGACCGACTAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 411
DB 370 CCTTCTGACCGACTAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
OY 412 GCTGAAAGAGCATGCGACCGCATTTATGAAATTCATCCAGACCTGAGACGAGGAGCG 471
DB 430 GCTGAAAGAGCATGCGACCGCATTTATGAAATTCATCCAGACCTGAGACGAGGAGCG 489
OY 472 CAGGAGGTGTGGAGCGGACAGTGAAGTGGCTGGGAGAGCCAGATCAACAGCTCCAGC 531
DB 490 CAGGAGGTGTGGAGCGGACAGTGAAGTGGCTGGGAGAGCCAGATCAACAGCTCCAGC 549
OY 532 CCTGCGCTACAAACCGGACGATCGGGAGTTTATGATCAACCGGCGCTGCTACTAC 591
DB 550 CCTGCGCTACAAACCGGACGATCGGGAGTTTATGATCAACCGGCGCTGCTACTAC 609
OY 592 CTGTACTGTCAAGTGACATTTGATGAGGGAAGGCTGTACTGAAAGCTGAGCTTGTG 651
DB 610 CTGTACTGTCAAGTGACATTTGATGAGGGAAGGCTGTACTGAAAGCTGAGCTTGTG 669
OY 652 GTGAGATGTGTGCTGCGGCTGCGGCTGCTGAGAGAAATTCACAGCCACTGCGGCAATTC 711
DB 670 GTGAGATGTGTGCTGCGGCTGCGGCTGCTGAGAGAAATTCACAGCCACTGCGGCAATTC 729
OY 712 CTGCGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 771

```

Db	730	CTGGGGCCCCAGCTCCGCTCTGGCAGGTGTGTGGCTGTTGGCCCTGGGGCCAGGGTCC	783
Oy	772	TCCTGGGATCCGACCCCTCCCTGGGGCCCACTCAAGGCTGCCCCCTTCTCTACCTAC	831
Db	790	TCCCTGGCGGATCCCGACACCTCTCCCTGGGGCCCACTCTCAAGGCTGCCCCCTTCTC	849
Oy	832	TTCCGACCTCTTCAGGTTCACTAGAGGGGCCCTG	865
Db	850	TTCCGACCTCTTCAGGTTCACTAGAGGGGCCCTG	883
RESULT 10			
ID	AAV18599	standard; cDNA; 1168 BP.	
AC	AAV18599;		
XX			
DT	21-JUL-1998	(first entry)	
XX			
DE	Mus musculus tumour necrosis factor related ligand (TRELL) gene.		
XX			
KW	TRELL; tumour necrosis factor related ligand; tnf; treatment;		
KW	cancer; autoimmune disease; immune system; stimulation; suppression;		
KW	graft rejection; ds.		
XX			
OS	Mus musculus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	2..679	
FT		/*tag= a	
FT		/note="tumour necrosis factor related ligand"	
XX			
PN	WO9805783-A1.		
PD	12-FEB-1998.		
XX			
PF	07-AUG-1997;	97WO-US13945.	
XX			
PR	18-MAR-1997;	97US-0040820.	
PR	07-AUG-1996;	96US-0023541.	
PR	18-OCT-1996;	96US-0028515.	
XX			
PA	(BIOU ) BIOGEN INC.		
PA	(UYGE-) UNIV GENEVA FACULTY MEDICINE.		
XX			
PI	Browning JL, Chicheportliche Y;		
DR	WPI: 1998-145619/13.		
DR	P-PSDB: AAW47524.		
XX			
PT	Tumour necrosis factor related ligand - useful for, e.g. treating		
PT	cancer, auto-immune disease and immune responses to tissue grafts		
XX			
PS	Claim 2; Pages 45-46; 69pp; English.		
XX			
CC	The sequence is that encoding mouse tumour necrosis factor related		
CC	ligand (TRELL). TRELL or active fragments can be included with a		
CC	carrier in pharmaceutical compositions to treat cancer, autoimmune		
CC	diseases or immune responses to tissue grafts, or to stimulate or		
CC	suppress the immune system. It is useful to screen for TRELL		
CC	receptors, by labelling with a detectable label and screening		
CC	compositions for binding. Agents interfering with TRELL-receptor		
CC	binding can also be screened for, can then be administered,		
CC	optionally with interferon- gamma, to induce cell death or		
CC	treat, suppress or alter immune responses (especially involving human		
CC	adenocarcinoma cells) involving a signal pathway between TRELL and its		
CC	receptor. The DNA sequence can be used in gene therapy for		
CC	TRELL-related disorders in mammals (especially humans), e.g. tumours,		
CC	autoimmune and inflammatory diseases or inherited genetic disorders,		
CC	by introducing into cells, and expressing, therapeutically effective		
CC	amounts of a vector, e.g. a virus comprising a gene encoding TRELL.		
CC	It may also be of use in the preparation of prepare probes for		
CC	screening natural/synthetic DNAs for TRELL-encoding sequences		

[illegible]

```

DB 937 CACCCACACGCTTTATGACTTGTGCAC----- 968
OY 1193 CACTTGACACTAAGAGGGGGCTGGACCTGGGGGAGGAGCCAAAGAGCTGGGCTAG 1252
DB 969 -----CAGCCTAGATGAGTGGCTGGACCTGGTGGAGAAAGCCAGAACTGGGACTAG 1023
OY 1253 GCCAGAGCTTCCCAATGTGAGGGGAGAG-AACAAGACAAAGCTCCCTCTGAGATTC 1311
DB 1024 GCCAGAAAGTTCCCACTGTGAGGGGAGAGAGCTGGGGACAAGCTCTCCCTGGA----TC 1079
OY 1312 CCTGTGATTTTAAACAGATTTATTTATTTATTTATTTGACAAATGTTGA 1366
DB 1080 CCTGTGATTTTGA--AGATACATATTTTATTTATTTATTTGACAAATGTTGA 1132

RESULT 11
AAK23425
ID AAK23425 standard; DNA; 701 BP.
XX
AC AAK23425;
XX
DT 18-JUN-1999 (first entry)
XX
DE Mouse TNRL3 DNA.
XX
KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KM developmental abnormality; gestational abnormality; prostate cancer;
KM APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KM apoptosis; mouse; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 1..636
FT /tag= a
FT /product= "TNRL3"
XX
XX WO9911791-A2.
XX
XX 11-MAR-1999.
XX
XX PD
XX
PF 04-SEP-1998; 98WO-US18393.
XX
XX PR 05-SEP-1997; 97US-0924634.
XX
XX (UNIV ) UNIV WASHINGTON.
XX
XX PA
XX PI Chaudhary PM;
XX
XX WPI; 1999-205191/17.
XX
XX DR P-PSDB; AAM93591.
XX
XX PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
XX PT useful for diagnosis and treatment of prostate cancer and
XX PT developmental or gestational abnormalities
XX
XX PS Example VII; Fig 13B; 156bp; English.
XX
XX CC This invention describes isolated Tumor Necrosis Factor (TNF) family
XX CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
XX CC their active fragments. APO4 is useful for diagnosing prostate cancer
XX CC by determining levels of APO4 in an individual. Prostate cancer can also
XX CC be treated using APO4 selective binding agents linked to a therapeutic
XX CC moiety. APO4 polypeptides are also useful for identifying selective
XX CC binding agents, useful in diagnosis/treatment of disease by binding of
XX CC agents to the polypeptide/active fragment which is extracellular, or
XX CC expressed on the cell surface. The binding is preferably performed in
XX CC vivo. APO4 polypeptides/active fragments are also useful for screening
XX CC for agonists and antagonists by binding and observing the changer in APO4
XX CC activity. Effective pharmacological agents useful in diagnosis or

```

```

CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
XX Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;
SO
Query Match 37.8%; Score 519.2; DB 20; Length 701;
Best Local Similarity 87.3%; Pred. No. 3.4e-101;
Matches 569; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
OY 220 CTGGCCGTGTCAGTTTGGGAGACCGGAGCATGCTGTCCCGCAGAGAGCTCCAGAG 279
DB 1 CTGTGTGTGTGTCAGCTTGAGGAGCTGGGAGAGCTGTCTGCCAGAGAGCTTCTCAGAG 60
OY 280 GAGCTGTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 339
DB 61 GAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
OY 340 CAGGATCTGCGCTTCTCTGAAACGACTAGTTGGGCTTCGAGAGAGTCACTAAAGGC 399
DB 121 CAGGATGTGTACCTTTCTTTGAAACACTAGTCCGCTCGAAGAGTCTCTAAAGGC 180
OY 400 CGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 459
DB 181 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
OY 460 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 519
DB 241 CAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
OY 520 AACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579
DB 301 AACAGCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 580 GGGCTCTACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639
DB 361 GGGCTCTACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
OY 640 CTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 699
DB 421 CTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
OY 700 GGGGCGAGTTCCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 759
DB 481 GCGAGAGAGCTCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
OY 760 CGGCGAGAGTCTCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
DB 541 CGGCGAGAGTCTCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
OY 820 TTCTCTCACTACTTGGAGCTTTCCAGGTTCACTGAGGGGCCCTGGCTGCC 871
DB 601 TTCTCACTACTTGGAGCTTTCCAGGTTCACTGAGGGGCCCTGGCTGCC 652

RESULT 12
AAK22190
ID AAK22190 standard; cDNA to mRNA; 282 BP.
XX
XX AAK22190;
XX
AC 27-AUG-1996 (first entry)
XX
XX DE Human gene signature HUMGS03761.
XX
XX KW Gene signature: messenger RNA; mRNA; relative abundance; frequency;
XX KW human; Cloning; mapping; non-biased library; diagnosis; detection;
XX KW cell typing; abnormal cell function; ss.

```





Search completed: March 31, 2003, 08:17:36  
Job time : 437.06 secs

PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;  
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
PI Lanson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
DR WPI: 1999-494092/41.

XX Novel human genes and their expression products which are  
PT differentially expressed in different cell types  
XX

PS Claim 1: Page 2250-2251: 2479pp; English.

XX The present invention describes a library of human polynucleotides  
CC comprising the sequences given in AA212532 to AA217779. Also described is  
CC a method of detecting differentially expressed genes correlated with the  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample from a cell  
CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.  
XX

SO Sequence 1337 BP: 42 A: 577 C: 27 G: 22 T: 669 other:

Query Match 4.4%; Score 61; DB 20; Length 1337;  
Best Local Similarity 32.1%; Pred. No. 0.0013;

Matches 145; Conservative 0; Mismatches 307; Indels 0; Gaps 0;

OY 690 CTCAGCCACTGGGCGGAGTTCCTCGGGCCCGAGCTCGGCTGTGGGCT 749  
DB 382 CANNNNCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 441  
OY 750 GTTGACCTCGAGGAGGCTCTCCGCGGATCCGACCTCCCTGGGCCATCTCAA 809  
DB 442 CCNCCGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 501  
OY 810 GGGTCCCCCTTCCTACCTACTTGGAGCTTTCAGGTTCACTGAGGGGCCCTGGTTC 869  
DB 502 NCNCCGCCGCCGCCGCCGCCNNCCNNCCNNCCNNCCNNCCNNCCNNCC 561  
OY 870 CCCACAGTGTCCAGAGGCTGCGGCTCCCTGAGAGCTCTGTGGGACCCGGTCCCTC 929  
DB 562 CCCNCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 621  
OY 930 TGCCCCACCTCAGCGGCTCTTGGTCCAGACCTCCCTCCCTAGAGGCTGCTGG 989  
DB 622 CCNCCNNCCGCCGCCGCCGCCNNNNNNNNNNNNNNNNNNNNNNNNNN 681  
OY 990 CCGTTCAGTGTTCATCCACATAAATACAGTATTCGACCTTATCTTACAACTC 1049  
DB 682 NNCNCCNNCCGCCGCCGCCGCCNNCCNNCCNNCCNNCCNNCCNNCC 741  
OY 1050 CCCCACGGCCACTCTCCAGCTAGTCCCAATCCCTGACCTTTGAGGCCGCCAG 1109  
DB 742 CCCCNCNCCGCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCC 801  
OY 1110 TGATCTGACTGCCCTGGCCAGACGCC 1141  
DB 802 NNNNNCCNNCCGCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCC 833

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:58:15 ; Search time 4329.73 Seconds

(without alignments)  
9228.782 Million cell updates/sec

Title: US-09-245-198A-3

Perfect score: 1373

Sequence: 1 atgcatctgttagacttga.....gacaaatgttgataaatg 1373

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: GenEmbl:*
2: gb_ba:*
3: gb_hcg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_pat:*
8: gb_ph:*
9: gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_sts:*
13: gb_sy:*
14: gb_un:*
15: gb_vl:*
16: em_ba:*
17: em_fun:*
18: em_hum:*
19: em_in:*
20: em_mu:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_un:*
30: em_vl:*
31: em_hcg_hum:*
32: em_hcg_inv:*
33: em_hcg_other:*
34: em_hcg_mus:*
35: em_hcg_pln:*
36: em_hcg_rtd:*
37: em_hcg_mam:*
38: em_hcg_vit:*
39: em_sy:*
40: em_hgtgo_hum:*
41: em_hgtgo_other:*
```

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320.2	96.2	1353	6 AX201324	AX201324 Sequence
2	1320.2	96.2	1368	9 AF055872	AF055872 Homo sapi
3	1285	93.6	1306	9 AF030099	AF030099 Homo sapi
4	1226.4	89.3	1236	6 AR140407	AR140407 Sequence
5	1096.8	79.9	1651	9 BC019047	BC019047 Homo sapi
6	768.2	56.0	177703	2 AC016876	AC016876 Homo sapi
7	754	54.9	215795	2 AC127470	AC127470 Pan trogl
8	629.2	45.8	898	6 AX180714	AX180714 Sequence
9	614.6	44.8	1168	10 AF030100	AF030100 Mus muscu
10	364.2	26.5	153553	2 AC126921	AC126921 Bos tauru
11	309.6	22.5	203083	2 AC069459	AC069459 Mus muscu
12	309.6	22.5	234182	10 AL603707	AL603707 Mouse DNA
13	254.2	18.5	138792	2 AC119115	AC119115 Rattus no
14	246.4	17.9	177555	2 AC130192	AC130192 Sus scrofa
15	242.4	17.7	161428	2 AC126825	AC126825 Canis fam
16	193.4	14.1	195	6 AX379024	AX379024 Sequence
17	158.6	11.6	184026	2 AC098923	AC098923 Rattus no
18	111	8.1	203281	2 AC126237	AC126237 Canis fam
19	100.6	7.3	139405	2 AC126239	AC126239 Felis cat
20	87.2	6.4	7218	6 I66494	I66494 Sequence 14
21	77.6	5.7	125020	9 AF429315	AF429315 Homo sapi
22	74.8	5.4	303091	2 AC084799	AC084799 Mus muscu
23	71	5.2	121451	2 AC111931	AC111931 Rattus no
24	70.8	5.2	73948	2 AC022556	AC022556 Homo sapi
25	70.6	5.1	144979	2 AC016280	AC016280 Homo sapi
26	70.4	5.1	220469	2 AC074307	AC074307 Mus muscu
27	70.2	5.1	298166	2 AC087563	AC087563 Homo sapi
28	69.2	5.0	936	8 CNS014YK	AL113556 Botrytis
29	69.2	5.0	82897	2 AC022595	AC022595 Homo sapi
30	68.2	5.0	110737	2 AC011105	AC011105 Homo sapi
31	68.2	5.0	265537	2 AC087228	AC087228 Mus muscu
32	68	5.0	181344	2 AC125880	AC125880 Rattus no
33	68	5.0	300695	2 AC079431	AC079431 Mus muscu
34	67.6	4.9	626	8 CNS018RP	AL110716 Botrytis
35	67.2	4.9	66624	2 AC101520	AC101520 Mus muscu
36	67.2	4.9	147124	2 AC022768	AC022768 Homo sapi
37	67.2	4.9	177883	2 AC046159	AC046159 Homo sapi
38	67.2	4.9	181988	2 AC090552	AC090552 Homo sapi
39	67	4.9	74138	2 AC021272	AC021272 Homo sapi
40	66.8	4.9	62649	2 AC022552	AC022552 Homo sapi
41	66.8	4.9	73628	2 AC021793	AC021793 Homo sapi
42	66.8	4.9	100511	2 AC010774	AC010774 Homo sapi
43	66.6	4.9	167830	2 AC103026	AC103026 Rattus no
44	66.4	4.8	840	8 CNS01BNS	AL114464 Botrytis
45	66.2	4.8	61834	2 AC068105	AC068105 Homo sapi

## ALIGNMENTS

```
RESULT 1
LOCUS AX201324 1353 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 3 from Patent WO0153486.
ACCESSION AX201324
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1353)
Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurey,A.L.,
Hillan,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
Stone,D.M., Watanabe,C.K. and Wood,W.I.
```



```

/product="Ap03/DR3 ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAARRSORRRGRREPEPTALLVPLALGLALACLGLLAVLSL
GSRASLSAEPDELVAEDQDPSSELPDPEESQDDPEFLNRLVPRSPKPKRTR
ARRIAAHVEHPRGDQDAGVAGVSVGSEARINSSPLRNROGEPITVRAGL
YYLCOVHFDEGKAVYLLKDLVDGVALRCLREFEASITAASSLQPLRLCOVSLAL
RPSLSRLRLPMALHNAKAPLTYEGLFOVH"
BASE COUNT      272 a      443 c      389 g      264 t
ORIGIN
Query Match      96.2%; Score 1320.2; DB 9; Length 1368;
Best Local Similarity 99.8%; Pred. No. 7.9e-231;
Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 49 CGATCCCTCCGGTCCCGGATGGGGGGCGGTGAGGACGACACCCCGCCCGCATG 108
DB 1 CGATCCCTCCGGTCCCGGATGGGGGGCGGTGAGGACGACACCCCGCCCGCATG 60
QY 109 GCCGCCCTCGAGCCAGAGGCGAGGGGCGCGGGGGAGACCGGCGACCCCTGCTG 168
DB 61 GCCGCCCTCGAGCCAGAGGCGAGGGGCGCGGGGGAGACCGGCGACCCCTGCTG 120
QY 169 GTCCCGCTCGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 228
DB 121 GTCCCGCTCGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 180
QY 229 GTCAATTTGGGGAGCGCGGCGATCGCTGTCGCCCGCAGAGCCCTGCCAGAGAGCTGGTG 288
DB 181 GTCAATTTGGGGAGCGCGGCGATCGCTGTCGCCCGCAGAGCCCTGCCAGAGAGCTGGTG 240
QY 289 GCAGAGAGAGACACAGACCCGTCGAGACTGAATCCCGACAGAGAGAGAGAGATCCT 348
DB 241 GCAGAGAGAGACACAGACCCGTCGAGACTGAATCCCGACAGAGAGAGAGATCCT 300
QY 349 GCGCCTTTCTGACCGACTAGTGGCCCTCGCAGAGTGCACCTAAAGCGCGAANA 408
DB 301 GCGCCTTTCTGACCGACTAGTGGCCCTCGCAGAGTGCACCTAAAGCGCGAANA 360
QY 409 CGGGCTCGAAGAGGATGGGACCCCATTTATGAAAGTTCAATCCAGACCTGGAGAGAGAG 468
DB 361 CGGGCTCGAAGAGGATGGGACCCCATTTATGAAAGTTCAATCCAGACCTGGAGAGAGAG 420
QY 469 GCGCAGGAGAGTGTGAGCGGAGACAGTGAAGTGGTGGGAGAGAGCAGATCAACAGCTCC 528
DB 421 GCGCAGGAGAGTGTGAGCGGAGAGACAGTGAAGTGGTGGGAGAGAGCAGATCAACAGCTCC 480
QY 529 AGCCTCTGCGCTACACCGCCAGATCGGGAGATTATAGTCACCCGGGCTGGGCTTAC 588
DB 481 AGCCTCTGCGCTACACCGCCAGATCGGGAGATTATAGTCACCCGGGCTGGGCTTAC 540
QY 589 TACCTGTACTGTACAGTCACTTTGATGAGGGGAGGGCTGTACTGTGAAGTGAAGTTG 648
DB 541 TACCTGTACTGTACAGTCACTTTGATGAGGGGAGGGCTGTACTGTGAAGTGAAGTTG 600
QY 649 CTGTGTGATGTGTGCTGGCCCTGCGCTGCGCTGAGAGAAATTTACGACCACTCGGCCAGT 708
DB 601 CTGTGTGATGTGTGCTGGCCCTGCGCTGCGCTGAGAGAAATTTACGACCACTCGGCCAGT 660
QY 709 TCCCTCGGGGCCACAGCTCGGCTCTGCGCAGAGTGTGAGGCTGTTGGCCCTGGGCCAGGG 768
DB 661 TCCCTCGGGGCCACAGCTCGGCTCTGCGCAGAGTGTGAGGCTGTTGGCCCTGGGCCAGGG 720
QY 769 TCCCTCGGGGCCACAGCTCGGCTCTGCGCAGAGTGTGAGGCTGTTGGCCCTGGGCCAGGG 828
DB 721 TCCCTCGGGGCCACAGCTCGGCTCTGCGCAGAGTGTGAGGCTGTTGGCCCTGGGCCAGGG 780
QY 829 TACTTCGAGACTTTCAGAGTTTACTGAGGGGCGCGTGTCCCAAGTGCATCCAGGCT 888
DB 781 TACTTCGAGACTTTCAGAGTTTACTGAGGGGCGCGTGTCCCAAGTGCATCCAGGCT 840
QY 889 GCGGCGCTCCCTCGACAGCTCTGTGGGCAACCCGCTGCCCTGTGCCCAACCTCAGCGCT 948

```

```

DB 841 GCGGCGCTCCCTCGACAGCTCTGTGGGCAACCGGCTGCCCTGTGCCCAACCTCAGCGCT 900
QY 949 CTTTCTCTCCAGACCTGCCCCCTTCTCTAGAGGCTCCCTGGGCTTTCACGTTTCCA 1008
DB 901 CTTTCTCTCCAGACCTGCCCCCTTCTCTAGAGGCTCCCTGGGCTTTCACGTTTCCA 960
QY 1009 TCCCAATTAATACAGTATTCACACTCTTATCTACACTCCCGCAGCGCCACTCTCCA 1068
DB 961 TCCCAATTAATACAGTATTCACACTCTTATCTACACTCCCGCAGCGCCACTCTCCA 1020
QY 1069 CCTCACTAGCTCCCAATTCCTGACCTTTTGAGGCCCCCAGTGAATCTGACTCCCCCTG 1128
DB 1021 CCTCACTAGCTCCCAATTCCTGACCTTTTGAGGCCCCCAGTGAATCTGACTCCCCCTG 1080
QY 1129 GCCACAGACCCCGAGGAGTGTGTCACTGTACTGTGTGGGCAAGAGTGGTCCAGAG 1188
DB 1081 GCCACAGACCCCGAGGAGTGTGTCACTGTACTGTGTGGGCAAGAGTGGTCCAGAG 1140
QY 1189 ACCCCACTTCAGAGCACTAAGAGGGGCTGACCTGGCGGCGAGAGGCCAAAGAGACTGGGC 1248
DB 1141 ACCCCACTTCAGAGCACTAAGAGGGGCTGACCTGGCGGCGAGAGGCCAAAGAGACTGGGC 1200
QY 1249 CTAGGCCAGAGAGTCCCAATGTAGAGGGCGAGAAACAAGACAGCTCCTCCTTGAGAA 1308
DB 1201 CTAGGCCAGAGAGTCCCAATGTAGAGGGCGAGAAACAAGACAGCTCCTCCTTGAGAA 1260
QY 1309 TTCCCTGTGATTTTAAACGATATATTTTATATATGAGCAAAATGTTGATA 1368
DB 1261 TTCCCTGTGATTTTAAACGATATATTTTATATATGAGCAAAATGTTGATA 1320
QY 1369 AATGG 1373
DB 1321 AATGG 1325

RESULT 3
AF030099 1306 bp mRNA linear PRI 20-DEC-1997
LOCUS AF030099
DEFINITION Homo sapiens TWEAK mRNA, complete cds.
ACCESSION AF030099
VERSION AF030099.1 GI:2707218
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1306)
Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,
Hession,C., Garcia,I. and Browning,J.L.
TWEAK, a new secreted ligand in the tumor necrosis factor family
that weakly induces apoptosis
J Biol. Chem. 272 (51), 32401-32410 (1997)
JOURNAL
MEDLINE 98070415
PUBMED 9405449
REFERENCE
2 (bases 1 to 1306)
Bourdon,P., Hession,C., Tizard,R. and Browning,J.
Direct Submission
JOURNAL
Submitted (14-Oct-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
source
1..1306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue-type="tonsil"
18..767
/note="Ligand in the TNF family; secreted protein; start
codon not verified experimentally"
/product="TWEAK"
/protein_id="AAC31923.1"
/db_xref="GI:2707219"

```

/translation="MAARRSQRNRGRBEPGTALLVPLALGLALACTLLAVSL  
GSRASLSAEPAQELVAEEDDPSLEINQTEESODPAFLRLIVPRBSAKRKT  
ARRAIAHYEPRGODAGVGTGMEPEARINSSPLRYNNOIEFLVTRAGL  
XYLGOVHDEBKAVYIKIDLLVDVLAIRCLIEEBSATMASSLGQLRLCOVSGLLAL  
RGGSSLRITTEPMHAKPAFLITTEGLFQVH"  
BASE COUNT 247 a 434 c 368 g 257 t  
ORIGIN  
Query Match 93.6%; Score 1285; DB 9; Length 1306;  
Best Local Similarity 100.0%; Pred. No. 2,1e-224;  
Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY CACAGCCCCCCCCCATGCGCCCGCTGCGAGCCAGAGCGCGAGGGGGCGCGGGGG 148  
DB 1 CACAGCGCCCCCGCCCATGCGCCCGCTGCGAGCCAGAGCGCGAGGGGGCGCGGGGG 60  
QY 149 AGCGGGGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 208  
DB 61 AGCGGGGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 209 TCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 268  
DB 121 TCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 269 CTGCGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 328  
DB 181 CTGCGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
QY 329 CAGAGAAAGCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 388  
DB 241 CAGAGAAAGCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
QY 389 CACCTAAAGCCGGAAGACGCGGCTGGAAGAGCGAGCTGCGAGCCATTATGAATTCTATC 448  
DB 301 CACCTAAAGCCGGAAGACGCGGCTGGAAGAGCGAGCTGCGAGCCATTATGAATTCTATC 360  
QY 449 CAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508  
DB 361 CAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
QY 509 AAGCCAGAAATCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568  
DB 421 AAGCCAGAAATCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 569 TCACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628  
DB 481 TCACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 629 TCTACCTGAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688  
DB 541 TCTACCTGAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 689 TCTCAGCAGCTGCGGCGAGTTCCTGCGGCGCCAGCTCCGCTCTGCGCAGAGTCTGCGG 748  
DB 601 TCTCAGCAGCTGCGGCGAGTTCCTGCGGCGCCAGCTCCGCTCTGCGCAGAGTCTGCGG 660  
QY 749 TGTGTGCGCTGCGGCGAGGCTCTCTCTGCGAGTTCGCGACCTCTCCCTGCGGCGCATCTCA 808  
DB 661 TGTGTGCGCTGCGGCGAGGCTCTCTCTGCGAGTTCGCGACCTCTCCCTGCGGCGCATCTCA 720  
QY 809 AGGCTGCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 868  
DB 721 AGGCTGCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
QY 869 CCCCACAGTGTGCCAGAGCTGCGGCGCTCCCTGAGACAGCTCTCTGCGCAGCCGCTGCTCT 928  
DB 781 CCCCACAGTGTGCCAGAGCTGCGGCGCTCCCTGAGACAGCTCTCTGCGCAGCCGCTGCTCT 840  
QY 929 CTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988  
DB 841 CTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 989 GCCTGTTCACGCTTTTCCATCCCATTAATACAGTATTCCACACTCTATTCTTACACT 1048

|||||  
DB 901 GCCTGTTCACGCTTTTCCATCCCATTAATACAGTATTCCACACTCTATTCTTACACT 960  
QY 1049 CCCCCACCGCCACCTGCTCCACCTCAGCTAGCTTCCCAATCCGAGCTTTGAGGGCCCCA 1108  
DB 961 CCCCCACCGCCACCTGCTCCACCTCAGCTAGCTTCCCAATCCGAGCTTTGAGGGCCCCA 1020  
QY 1109 GTGATCTGCATCTCCCTGCGCCAGACCCCGAGGAGATTGTCTCAGTACTCTGTG 1168  
DB 1021 GTGATCTGCATCTCCCTGCGCCAGACCCCGAGGAGATTGTCTCAGTACTCTGTG 1080  
QY 1169 GGCAGAGATGCTGCAAGAACCCCACTTCAGGCTAAGAGGGCTGAGACTGCGCGCA 1228  
DB 1081 GGCAGAGATGCTGCAAGAACCCCACTTCAGGCTAAGAGGGCTGAGACTGCGCGCA 1140  
QY 1229 GGAAGCCAAAGACAGTGGGCTAGGCGAGAGTTCGCAATGTGAGGGCGGAAACAG 1288  
DB 1141 GGAAGCCAAAGACAGTGGGCTAGGCGAGAGTTCGCAATGTGAGGGCGGAAACAG 1200  
QY 1289 ACAAGCTCTCTCTGAGATTCCTGAGATTTTAAACAGATATTATTTATTTATTT 1348  
DB 1201 ACAAGCTCTCTCTGAGATTCCTGAGATTTTAAACAGATATTATTTATTTATTTATTT 1260  
QY 1349 ATTGTGACAAATGTTGATTAATGG 1373  
DB 1261 ATTGTGACAAATGTTGATTAATGG 1285  
RESULT 4  
LOCUS AR140407 1236 bp DNA linear PAT 16-JUN-2001  
DEFINITION Sequence 1 from patent US 6207642.  
ACCESSION AR140407  
VERSION AR140407.1 GI:14482903  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1236)  
AUTHORS Wiley,S.R.  
TITLE Member of the TNF family useful for treatment and diagnosis of  
disease  
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;  
FEATURES  
source location/Qualifiers  
BASE COUNT 225 a 416 c 358 g 237 t  
ORIGIN  
Query Match 89.3%; Score 1226.4; DB 6; Length 1236;  
Best Local Similarity 99.5%; Pred. No. 1e-213;  
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 106 ATGGCGGCGCTGCGGAGCCAGAGGAGGAGGCGCGGGGAGGCGGGGACCGGCGCTG 165  
DB 1 ATGGCGGCGCTGCGGAGCCAGAGGAGGAGGCGCGGGGAGGCGGGGACCGGCGCTG 60  
QY 166 CTGCTCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 225  
DB 61 CTGCTCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 120  
QY 226 GTGCTCAGTTTGGGAGAGCGGCGATGCTGCTGCGGCGAGAGCTTCCAGAGAGAGCTG 285  
DB 121 GTGCTCAGTTTGGGAGAGCGGCGATGCTGCTGCGGCGAGAGCTTCCAGAGAGAGCTG 180  
QY 286 GTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 345  
DB 181 GTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240  
QY 346 CTGCGGCTTCTCTGAAACGACTAGTTCGCGCTCGGAGAGAGTGCACCTAAAGGCGGAAA 405  
DB 241 CTGCGGCTTCTCTGAAACGACTAGTTCGCGCTCGGAGAGAGTGCACCTAAAGGCGGAAA 300

```

Oy 406 ACACGGGCTCGAAGAGCGATCGACGCCCATATGAGTTTCATCCAGACCTGGACAGAC 465
    |||||||
Db 301 ACAGGGGCTCGAAGAGCGATCGACGCCCATATGAGTTTCATCCAGACCTGGACAGAC 360
Oy 466 GAGGCGCAGGAGGAGTGTGGAGCGGACAGTGAAGTGGCTGGAGGAAAGCCAGAAATCAACAC 525
    |||||||
Db 361 GAGGCGCAGGAGGAGTGTGGAGCGGACAGTGAAGTGGCTGGAGGAAAGCCAGAAATCAACAC 420
Oy 526 TCCAGACCCCTCGGCTACACCGCCAGATCGGGGAGTTTATAGTACCCCGGGCTGGGCTC 585
    |||||||
Db 421 TCCAGACCCCTCGGCTACACCGCCAGATCGGGGAGTTTATAGTACCCCGGGCTGGGCTC 480
Oy 586 TACTACTGTACTGTACAGTGCATTTGATGAGGGGAAAGGCTGTGTACTGTAGAGTGCAG 645
    |||||||
Db 481 TACTACTGTACTGTACAGTGCATTTGATGAGGGGAAAGGCTGTGTACTGTAGAGTGCAG 540
Oy 646 TTGCTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 705
    |||||||
Db 541 TTGCTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Oy 706 ACTTCCTCGGGGCGCCAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 765
    |||||||
Db 601 ACTTCCTCGGGGCGCCAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGG 660
Oy 766 GGGTCTCTCTCGGAGATCGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 825
    |||||||
Db 661 GGGTCTCTCTCGGAGATCGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Oy 826 ACCTACTTGTGAGCTTTCCAGGTTTCTAGTGAAGGCGCTGTGTGTGTGTGTGTGTGTGT 885
    |||||||
Db 721 ACCTACTTGTGAGCTTTCCAGGTTTCTAGTGAAGGCGCTGTGTGTGTGTGTGTGTGTGT 780
Oy 886 GTGCGGGGCTCGGCTCGGACGCTCTGTGGGACCGGCTCGGCTCGGCTCGGCTCGGCTCGG 945
    |||||||
Db 781 GTGCGGGGCTCGGCTCGGACGCTCTGTGGGACCGGCTCGGCTCGGCTCGGCTCGGCTCGG 840
Oy 946 GCTCTTGTGTCCAGACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1005
    |||||||
Db 841 GCTCTTGTGTCCAGACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Oy 1006 CCATCCCAATTAATACAGTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1065
    |||||||
Db 901 CCATCCCAATTAATACAGTATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Oy 1066 CCACCTACTAGCTCCCAATCCCTGAGCCCTTGAAGGCGCCAGTGCATCTCCCTCC 1125
    |||||||
Db 961 CCACCTACTAGCTCCCAATCCCTGAGCCCTTGAAGGCGCCAGTGCATCTCCCTCCCTCC 1020
Oy 1126 CTGGCCAGACAGCCCGCAGGCGATGTGTCTACTGTACTGTGTGGCAAGATGGGTCCAG 1185
    |||||||
Db 1021 CTGGCCAGACAGCCCGCAGGCGATGTGTGTCTACTGTACTGTGTGGCAAGATGGGTCCAG 1080
Oy 1186 AAGACCCCACTTCAGGACCTAAGAGGGGCTGAGCTGCGGCGAGAGCAAGCCAAAGACTG 1245
    |||||||
Db 1081 AAGACCCCACTTCAGGACCTAAGAGGGGCTGAGCTGCGGCGAGAGCAAGCCAAAGACTG 1140
Oy 1246 GGGCTAGGCGAGAGTTCCTCCAAATGTGAGGGCGGAGAAACAGACAGAGTCTCTCCCTGA 1305
    |||||||
Db 1141 GGGCTAGGCGAGAGTTCCTCCAAATGTGAGGGCGGAGAAACAGACAGAGTCTCTCCCTGA 1200
Oy 1306 GAATTCCTGTGATTTTAAACAGATATTTATTTT 1341
    |||||||
Db 1201 GAATTCCTGTGATTTTAAACAGATATTTATTTT 1236

```

```

RESULT 5
BC019047      1651 bp  mRNA  linear  PRI 11-DEC-2001
LOCUS         Homo sapiens, similar to tumor necrosis factor (ligand)
DEFINITION    superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
ACCESSION     BC019047
VERSION       BC019047.1 GI:17512138

```

**KEYWORDS** MGC.  
**SOURCE** Homo sapiens.  
**ORGANISM** Homo sapiens.  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1651)  
**AUTHORS** Strausberg, R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-7590, USA  
**REMARK** NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
**COMMENT** Contact: MGC help desk  
Email: [cgaps@remail.nih.gov](mailto:cgaps@remail.nih.gov)  
Tissue Procurement: Louis Staudt  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (ILNL)  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc-mgc@nhgri.nih.gov](mailto:nisc-mgc@nhgri.nih.gov)  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J., Tlonsong, E.E., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 30 Row: P Column: 5  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596.

#### FEATURES

source

1..1651  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:20669 IMAGE:4766071"  
/tissue\_type="Primary B-cells from Tonsils"  
/clone\_lib="NIH-MGC-48"  
/lab\_host="DH10B-R"  
/note="vector: pOTB7"  
106..510  
/codon\_start=1  
/product="Similar to tumor necrosis factor (ligand)  
superfamily, member 12"  
/db\_xref="GI:17512139"  
/db\_xref="GI:17512139"  
/translation="MAARSSQRRRGRRGEPSTALLVPLALGLALACGLLAVSL  
GSRASLSAQPDAQELVAEDQDPSELNPOEEQDPAFLNLRVRRPRAKPKRKR  
ARRAIAHYEVHPPGODGAQADGQYTCCLRP"  
BASE COUNT 344 a 517 c 481 g 309 t  
ORIGIN

Query Match 79.9%; Score 1096.8; DB 9; Length 1651;  
Best Local Similarity 84.8%; Pred. No. 4.1e-190;  
Matches 1347; Conservative 0; Mismatches 2; Indels 239; Gaps 1;

```

Oy 25 TCCGGCGCGGCGCTCCCTCCCGATCCCTCGGCTCCCGGATGAGGGGCGGTGAGG 84
    |||||||
Db 25 TCCGGCGCGGCGCTCCCTCCCGATCCCTCGGCTCCCGGATGAGGGGCGGTGAGG 84
Oy 85 CAGGACAGCCCGCCCGCCATGCGCGCGCTCGGAGCAGAGCGAGGCGCGCGG 144
    |||||||
Db 85 CAGGACAGCCCGCCCGCCCGCCATGCGCGCGCTCGGAGCAGAGCGAGGCGCGCGG 144
Oy 145 GGGAGCGCGGAGCAGCGCTGCTGCTCGCGCTGCGGCTGAGGCTGAGGCTGAGG 204
    |||||||
Db 145 GGGAGCGCGGAGCAGCGCTGCTGCTCGCGCTGCGGCTGAGGCTGAGGCTGAGG 204

```

OY	205	TGCTCGGGCCCTCGTCGTCGGCCGTGGTCAGTTTGGGAGCGGGCAATCGCTGTGCCCCAG	264
Db	205	TGCTCGGGCCCTCGTCGTCGGCCGTGGTCAGTTTGGGAGCGGGCAATCGCTGTGCCCCAG	264
OY	265	GAGCCTGCCAGGAGGAGCTGTGGCGAGAGGAGACAGAGACCCGTGCGAATCGATCC	324
Db	265	GAGCCTGCCAGGAGGAGCTGTGGCGAGAGGAGACAGAGACCCGTGCGAATCGATCC	324
OY	325	CAGACAGAGAAAGCCAGAGATCCTGGCCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA	384
Db	325	CAGACAGAGAAAGCCAGAGATCCTGGCCCTTTCCTGAACCGACTAGTTCGGCCTCGCAGA	384
OY	385	AGTGACCTTAAGGCGGAAAAACACGGGGCTGGAAGAGCGATGGACCCCATATGAAGT	444
Db	385	AGTGACCTTAAGGCGGAAAAACACGGGGCTGGAAGAGCGATGGACCCCATATGAAGT	444
OY	445	CATCCACGACCTTGACAGAGACGAGCGACAG-----	475
Db	445	CATCCACGACCTTGACAGAGACGAGCGACAG-----	475
OY	476	-----	475
Db	505	CCATGAGATCTAAGTGTGGAGCCAGATTGTAACCCAGCTAGATGTGCTGTACT	564
OY	476	-----	475
Db	555	GGACATGCTCTGATGAAAGGCCAGGTGCTGCTGCAGGGGTGAGGGGTCATTCAGAGGG	624
OY	476	-----	475
Db	625	CCACATCCAAAAGGGGAGAGGGAGATTTCAGAAAAGAGAGAGACATCTCCACACA	684
OY	476	-----CAGCTGTGACGGGACAGTGAATGCTGGG	505
Db	685	TTACAGAGGGTCAAAGGAGAGAACAGAGACCAAGTGTGGACGGGAGACGTGCTGGG	744
OY	506	AGGAAGCCAGAAATACAGAGCTCCAGCCCTGTGGGTGTAACAACCGCCAGATCGGGAGATT	565
Db	745	AGGAAGCCAGAAATACAGAGCTCCAGCCCTGTGGGTGTAACAACCGCCAGATCGGGAGATT	804
OY	566	TAGTCACCCGGGCTGGGCTGTACTACTGTACTGTCTCAGGTGCACTTTGATGAGGGGAAG	625
Db	805	TAGTCACCCGGGCTGGGCTGTACTACTGTCTCAGGTGCACTTTGATGAGGGGAAG	864
OY	626	CTGTCTACTCTAAGCTGAGCTTGTGTGTGATGTGTGCTGGCCCTGCGCTCTGTAGAG	685
Db	865	CTGTCTACTCTAAGCTGAGCTTGTGTGTGATGTGTGCTGGCCCTGCGCTCTGTAGAG	924
OY	686	AATTCCTCAGCACTGCGGCGCATGTCCCTGGGGCCCGAGCTCGCGCTCTCAGAGGTCTG	745
Db	925	AATTCCTCAGCACTGCGGCGCATGTCCCTGGGGCCCGAGCTCGCGCTCTCAGAGGTCTG	984
OY	746	GCGTGTGGCCCTCGGCGCAGGGGTCTCCCTGCGGATCCGCAACCCCTCGGGGCCATC	805
Db	985	GCGTGTGGCCCTCGGCGCAGGGGTCTCCCTGCGGATCCGCAACCCCTCGGGGCCATC	1044
OY	806	TCAAGGCTGCCCCCTTCTCTACCTACTTGGGACTTTTCCAGGTTCACTGAGGGGCCCTG	865
Db	1045	TCAAGGCTGCCCCCTTCTCTACCTACTTGGGACTTTTCCAGGTTCACTGAGGGGCCCTG	1104
OY	866	TCTCCCACTGCTGCCAGGCTGCGGGCTCCCTGAGACAGTCTGTGGGCACCGGTC	925
Db	1105	TCTCCCACTGCTGCCAGGCTGCGGGCTCCCTGAGACAGTCTGTGTGGGCACCGGTC	1164
OY	926	CCCTGCGCCCACTTCAGCGCGCTTGTGTCAGAGACTGCGCCCTCCCTCTTAAGGCTGC	985
Db	1165	CCCTGCGCCCACTTCAGCGCGCTTGTGTCAGAGACTGCGCCCTCCCTCTTAAGGCTGC	1224
OY	986	TGGGCTGTTCACGTGTTTTTCATCCACATTAATACATATTCCACACTTATCTTACA	1045
Db	1225	TGGGCTGTTCACGTGTTTTTCATCCACATTAATACATATTCCACACTTATCTTACA	1288
OY	1046	ACTCCCCACCGCCACTCTTCACCTTACAGTACCTGCCCAATCCGACCCCTTTGAGGCC	1105

Db	1285	ACTCCCCACCGCCACACTCTCCACTCACTAGCTCTCCAAATCCCTGACCCCTTTGAGGCC	1344
Qy	1106	CCAGTGATCTGAGACTCCCCCCTGGCCACAGACCCCCAGAGGCAATTGTCTCACTGACTCT	1165
Db	1345	CCAGTGATCTGAGACTCCCCCCTGGCCACAGACCCCCAGAGGCAATTGTCTCACTGACTCT	1404
Qy	1166	GTTGGCCAAAGGATGGGTCCAGAAAGACCCCACTTCAGGCACTAAAGAGGGCTGGACTGGCG	1225
Db	1405	GTTGGCCAAAGGATGGGTCCAGAAAGACCCCACTTCAGGCACTAAAGAGGGCTGGACTGGCG	1464
Qy	1226	GCAGGAAGCCAAAGAGAGCTGGGGCTAGGCCCAAGAGATTTCCCAATATGTGAGGGCGAGAAC	1285
Db	1465	GCAGGAAGCCAAAGAGAGATGGGGCTAGGCCCAAGAGATTTCCCAATATGTGAGGGCGAGAAC	1524
Qy	1286	AAGACAAGCTCTCCCTTGTAGAAATTCCTGTGGATTTTAAACAGATATTAATTTATT	1345
Db	1525	AAGACAAGCTCTCCCTTGTAGAAATTCCTGTGGATTTTAAACAGATATTAATTTATT	1584
Qy	1346	ATTATTTGTACAAATGTTGATTAATGG	1373
Db	1585	ATTATTTGTACAAATGTTGATTAATGG	1612

	RESULT_6				
	AC016876/c				
	LOCUS				
	DEFINITION	AC016876 Homo sapiens clone RP11-16B87, *** SEQUENCING IN PROGRESS ***, 10			
	ACCESION	AC016876			
	VERSION	AC016876.5 GI:21313830			
	KEYWORDS	HGT; HIGS_PHASEI; HIGS_FULLTOP; HIGS_ACTIVEFIN.			
	SOURCE	human.			
	ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Britten,B., Nusbaum,C. and Lander,E. 1 (bases 1 to 177703)				
AUTHORS	Homo saplens chromosome, clone RP11-16B87 Unpublished				
TITLE	2 (bases 1 to 177703)				
JOURNAL	Britten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Balderin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collimore,A., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L., Howland,J.C., Johnson.R., Jones.C., Kann,L., Karkats,A., Klein,J., Lehoccky,J., Liu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,F., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J., Morrow,J., Naylor,T., Norman,C.H., O'Connor.T., O'Donnell.P., Peterson.K., Pollara,V., Riley,R., Roy.A., Santos,R., Severy,P., Stange-Thomann.N., Stojanovic.N., Subramanian,A., Telams.J., Tefaye.S., Tirrell,A., Vassiliev,H., Vo.A., Wheeler,J., Wu,X., Wyman.D., Ye.W.J., Zimmer.A. and Zody.M.				
DIRECT SUBMISSION					
TITLE	Submitted (08-Dec-1999) Whitehead Institute/MIT Center for Genome Researchch, 320 Charles street, Cambridge, MA 02141, USA .				
JOURNAL	3 (bases 1 to 177703)				
REFERENCE	Britten,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhalter,B., Camarta,J., Chang,J., Chazarro.B., Choepel,Y., Collimore,A., Cook,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faró,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham.L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones.C., Kanat,A., Karatas,A., Kelly,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu.G., Maclean,C., MacDonald,P., Major,J., Matthews.C., McCCarthy.M., Melidrim,J., Menes.L., Milhovat,T., Mlenka.V., Murphy,T., Naylor,J., Nguyen.C., Nicol.R., Norbu,C., Norman.C.H., O'Connor.T., O'Donnell.P., O'Neill,D., Oliver.J., Peterson.K., PhunKhang.P., Pierre,N., Raymond.C., Retta,R., Risse.C., Rogov,P., Roman,J., Roy.A., Schauer,R., Sculpback,R., Seamans.S., Severy.P.,				



```
Web site: http://www.nisc.nih.gov
Contact: nisc-zoehng@nsl.nih.gov
----- Project Information -----
Center project name: cms
Center clone name: 145D13
----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203929 bases at least Q40
Consensus quality: 206865 bases at least Q30
Consensus quality: 209181 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 214395; sum-of-ctrls
Quality coverage: 7.30x in Q20 bases; agarose-fp
Quality coverage: 6.47x in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 3593: contig of 3593 bp in length
*
3594 3693: gap of unknown length
*
3694 8001: contig of 4308 bp in length
*
8002 8101: gap of unknown length
*
8102 15639: contig of 7538 bp in length
*
15640 15739: gap of unknown length
*
15740 24775: contig of 9036 bp in length
*
24776 24875: gap of unknown length
*
24876 30150: contig of 5275 bp in length
*
30151 30250: gap of unknown length
*
30251 39212: contig of 8962 bp in length
*
39213 39312: gap of unknown length
*
39313 49342: contig of 10030 bp in length
*
49343 49442: gap of unknown length
*
49443 60956: contig of 11514 bp in length
*
60957 61056: gap of unknown length
*
61057 74520: contig of 13464 bp in length
*
74521 74620: gap of unknown length
*
74621 89522: contig of 14902 bp in length
*
89523 89622: gap of unknown length
*
89623 105537: contig of 15915 bp in length
*
105538 105637: gap of unknown length
*
105638 121197: contig of 15560 bp in length
*
121198 121287: gap of unknown length
*
121298 138883: contig of 17586 bp in length
*
138884 138983: gap of unknown length
*
138984 166802: contig of 27819 bp in length
*
166803 215795: gap of unknown length
*
166903 215795: contig of 48893 bp in length.
FEATURES
source
1. 215795
Location/Qualifiers
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone_lib="RP43-145D13"
/clone_lib="RP43"
1. 3593
/misc_feature
/note="assembly-fragment"
3694. 8001
/note="assembly-fragment"
8102. 15639
/note="assembly-fragment"
15740. 24775
/note="assembly-fragment"
24876. 30150
/note="assembly-fragment"
30251. 39212
/note="assembly-fragment"
39313. 49342
/misc_feature
/note="assembly-fragment"
49443. 60956
/note="assembly-fragment"
61057. 74520
/note="assembly-fragment"
74621. 89522
/note="assembly-fragment"
89623. 105537
/note="assembly-fragment"
105638. 121197
/note="assembly-fragment"
121298. 138883
/note="assembly-fragment"
138984. 166802
/note="assembly-fragment"
166903. 215795
/note="assembly-fragment"
BASE COUNT 52160 a 54859 c 53752 g 55513 t 1511 others
ORIGIN
Query Match 54.9%; Score 754; DB 2; Length 215795;
Best Local Similarity 99.2%; Pred. No. 9,1e-128;
Matches 768; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
misc_feature /note="assembly-fragment"
49443. 60956
/note="assembly-fragment"
61057. 74520
/note="assembly-fragment"
74621. 89522
/note="assembly-fragment"
89623. 105537
/note="assembly-fragment"
105638. 121197
/note="assembly-fragment"
121298. 138883
/note="assembly-fragment"
138984. 166802
/note="assembly-fragment"
166903. 215795
/note="assembly-fragment"
misc_feature
1260 GTTCCCAAAATGTGAGGGGAGAGAAAGCAAGACTCTCCCTTGGAATTCCTGTGGA 1319
|||||
Db 200116 GTTCCCAAAATGTGAGGGGAGAGAAAGCAAGACTCTCCCTTGGAATTCCTGTGGA 200115
|||||
Qy 1320 TTTTAAACAGATATTATTATTATTATTATTATTGTGACAAAATGTGATAATGG 1373
|||||
```



|||||  
DB 200176 TTTTAAACACATATATTTATTTATTTATGTCAGCAAGCTGTGATTAATGG 200229

RESULT 8  
AX180714 898 bp DNA linear PAT 06-AUG-2001  
LOCUS AX180714  
DEFINITION Sequence 1 from Patent WO0145730.  
ACCESSION AX180714  
VERSION AX180714.1 GI:15132570  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct.  
artificial sequences.  
REFERENCE  
1 (bases 1 to 898)  
AUTHORS  
TITLE  
JOURNAL  
Wiley, S.R.  
Patent: WO 0145730-A 1 28-JUN-2001;  
IMMUNEX CORPORATION (US)  
FEATURES  
Location/Qualifiers  
1..898  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
52..873  
/note="human TWEAK fusion protein construct"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC50004.1"  
/db\_xref="GI:15132571"  
/translation="MATGSRSTSLLAFLGLCLPMLQEGSSATSDRMKOIEDKIEEILS  
KIHVENEIARIKILIGERTSSLSRASLSAEPDAEELAEEDPDSEINPOTEEES  
ODPAPFLNRILVRPRRSAPKGRKTRARAIAMHVEVHPRGDQAGVDSVSGMEEA  
RINSSSPRLRVROIGERTVTRAGLYLYLCOVHPDEGRKRVYLKLDLVNGLALRLLEE  
FSATRASSLGLQLHCOVSGILALRPPSSLNIRLTPMHLKAAPLITFGLFQVH"

BASE COUNT 187 a 266 c 267 g 178 t  
ORIGIN

Query Match 45.8%; Score 629.2; DB 6; Length 898;  
Best Local Similarity 99.5%; Pred. No. 7; 6e-105;  
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 232 AGTTGGGAGCGGGGCATCGCTGCCGCCAGAGACCTGCCCAGAGAGAGAGCTGTGCA 291  
|||||  
DB 250 AGTTGGGAGCGGGGCATCGCTGCCGCCAGAGACCTGCCCAGAGAGAGAGCTGTGCA 309  
QY 292 GAGGAGAGACAGACCCCGTCGGAAGTGAATCCCGACAGAGAGAGAGAGAGATCTGCG 351  
|||||  
DB 310 GAGGAGAGACAGACCCCGTCGGAAGTGAATCCCGACAGAGAGAGAGAGAGATCTGCG 369  
QY 352 CTTTCCCTGAACCGAGTATGTCGCTCGCAGAGTGCACCTAAAGCGCGGAAACACAGG 411  
|||||  
DB 370 CTTTCCCTGAACCGAGTATGTCGCTCGCAGAGTGCACCTAAAGCGCGGAAACACAGG 429  
QY 412 GCTGAGAGAGCGATGCGACCCCATTTATGAAGTTTCATCCACGACTGAGACGAGAGCG 471  
|||||  
DB 430 GCTGAGAGAGCGATGCGACCCCATTTATGAAGTTTCATCCACGACTGAGACGAGAGCG 489  
QY 472 CAGGAGAGTGTGAGCGGAGCATGAGTGGCTGGGAGAGAGAGAGAGAGAGATCAACAGCTCCACG 531  
|||||  
DB 490 CAGGAGAGTGTGAGCGGAGCATGAGTGGCTGGGAGAGAGAGAGAGAGATCAACAGCTCCACG 549  
QY 532 CTTGCGCCTTACAAACCGCGAGATCGGGAGATTATAGTACCCGGGCTGGGCTTACTACT 591  
|||||  
DB 550 CTTGCGCCTTACAAACCGCGAGATCGGGAGATTATAGTACCCGGGCTGGGCTTACTACTAC 609  
QY 592 CTGTAAGTGTCAAGTCACTTTGATGAGGGAGAGGCTGTCTTACCTGAAGCTGAGACTTGCTG 651  
|||||  
DB 610 CTGTAAGTGTCAAGTCACTTTGATGAGGGAGAGGCTGTCTTACCTGAAGCTGAGACTTGCTG 669  
QY 652 GTGATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711  
|||||  
DB 670 GTGATGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729

QY 712 CTCGGGCCCCAGCTCCGCTGCGCAGGCTGTGGGCTGTGGCCCTGCGGCGAGGCTCC 771  
|||||  
DB 730 CTCGGGCCCCAGCTCCGCTGCGCAGGCTGTGGGCTGTGGCCCTGCGGCGAGGCTCC 789  
QY 772 TCCCTGCGGATCCGACACCTCCCTGCGGCGCATATCAAGCTGACCCCTTCTCTACCTAC 831  
|||||  
DB 790 TCCCTGCGGATCCGACACCTCCCTGCGGCGCATATCAAGCTGACCCCTTCTCTACCTAC 849  
QY 832 TTGCGACTCTTCAGGTTCACTGAGGGGCGCTGCG 865  
|||||  
DB 850 TTGCGACTCTTCAGGTTCACTGAGGGGCGCTGCG 883

RESULT 9  
AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
LOCUS AF030100  
DEFINITION Mus musculus TWEAK mRNA, partial cds.  
ACCESSION AF030100  
VERSION AF030100.1 GI:2707220  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 1168)  
Hession, C., Garcia, I. and Browning, J.L.  
TWEAK, a new secreted ligand in the tumor necrosis factor family  
that weakly induces apoptosis  
J. Biol. Chem. 272 (51), 32401-32410 (1997)  
JOURNAL  
MEDLINE  
PUBMED  
9405449  
REFERENCE  
2 (bases 1 to 1168)  
Chicheportiche, Y., Bixler, S., Tizard, R. and Browning, J.  
Direct Submission  
Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,  
Cambridge, MA 02142, USA  
JOURNAL

FEATURES  
Location/Qualifiers  
1..1168  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/cell\_type="peritoneal macrophages"  
<1..679  
/note="ligand in the TNF family; secreted protein"  
/codon\_start=2  
/product="TWEAK"  
/protein\_id="AAC53517.1"  
/db\_xref="GI:2707221"  
/translation="VLSGLALACLLLVVSLGSNATLSAOPSPDELTAEDEREP  
PELNQTEESDDVYFELQVLRPRRSAPKGRKARARARAIAMHVEVHPRGDQAGV  
DGYVSGWEETKINSSPRLRVROIGERTVTRAGLYLYLCOVHPDEGRKRVYLKLDLVN  
GLVLRCLLEERSATRASSPGLRQVSGILALRPPSSLNIRLTPMHLKAAPLITF  
GLFQVH"

BASE COUNT 242 a 360 c 298 g 268 t  
ORIGIN

Query Match 44.8%; Score 614.6; DB 10; Length 1168;  
Best Local Similarity 75.8%; Pred. No. 3; 4e-102;  
Matches 906; Conservative 0; Mismatches 219; Indels 70; Gaps 9;

QY 178 GCGTGGGCGCTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 237  
|||||  
DB 2 GGTCTGAGCGCTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 61  
QY 238 GGGAGCGGGCATCGCTGTCGCCAGAGAGCTGCCAGAGAGAGCTGTGCGAGAGAG 297  
|||||  
DB 62 GGGAGCTGGGCAACCTGCTGCTGCCAGAGAGCTTTCAGAGAGAGCTGACAGCAGAGAG 121  
QY 298 GACCAAGACCCCTCGGAATGAATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGATCTGCGCTTTC 357  
|||||  
DB 122 CGCGGGAGCCCTGGAATGAATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGATCTGCTTTC 181  
QY 358 CTGAACCGACTGTCGCTGCGCTGCGAGAGTGCACCTAAAGCGCGGAAACACGAGGCTCGA 417

```

Db      182 TTGGAACAACCTAGTCCGGCTCCGAAAGAGCTCTAAAGGCGGAGGCGGCGCTCCG 241
Oy      418 AGAGCGATCGACGCCCATTTATGAAATTATCCAGACCTGGACAGGACGGCAGGCA 477
Db      242 CAGAGCTATTGCAAGCCCATTTATGAGTTTCACTCCGCGCAGACAGATGAGACACA 301
Oy      478 GGTGTGACGGGACAGTGTGAGTGGTGGAGAGAACCAAGATCAAGAGCTCCAGCCCTG 537
Db      302 GGTGTGATGGAGACAGTGTGAGTGGTGGAGAGAACCAAGATCAAGAGCTCCAGCCCTG 361
Oy      538 CGGTAAACCCGCAAGATTCGGGGAGTTTATAGTACCCGGGCTGAGCTTACTACTG 597
Db      362 CGGTAGACCGCCAGATTCGGGGAGTTTATAGTACCCGGGCTGAGCTTACTACTG 421
Oy      598 TGTGAGTGCACACTTGTATGAGGAGGAGGCTGTACTGAACTGAGCTTGGCTGAG 657
Db      422 TGTGAGTGCACACTTGTATGAGGAGGAGGCTGTACTGAACTGAGCTTGGCTGAG 481
Oy      658 GGTGTGCTGGCCCTGGCGCTGCGTGAAGAAATTCACGCCACTGCGGCGCACTTCG 717
Db      482 GGTGTGCTGGCCCTGGCGCTGCGTGAAGAAATTCACGCCACTGCGGCGCACTTCG 541
Oy      718 CCCGACCTCGCCTCTGCCAGAGTGTGTGGCTTGGCCCTGGCGGAGGCTTCCTCG 777
Db      542 CCCGACCTCGCCTTGTGCGCAGAGTGTGTGGCTTGGCCCTGGCGGAGGCTTCCTCG 601
Oy      778 CGGATCCGACCCCTGCCCTGGCGGCGCAATTCACGCCACTGCGGCGCACTTCG 837
Db      602 CGGATCCGACCCCTGCCCTGGCGGCGCAATTCACGCCACTGCGGCGCACTTCG 661
Oy      838 CTCTTCAGAGTTCACTGAGGAGGCGCTGTCTCCCAAGTGTGTCCAGAGCTGCGGCTC 897
Db      662 CTCTTCAGAGTTCACTGAGGAGGCGCTGTCTCCCAAGTGTGTCCAGAGCTGCGGCTC 721
Oy      898 -----CCTGACAGCTCTGTGGGACCCGCTGCCCTGTGCGGCGCACTTCG 952
Db      722 AGGAGATCAGCAGACCTCCCTGAGAGGCGCTGTGCGGCGCACTTCG 780
Oy      953 GCTCAGACCTGCGCCTCCCTGAGAGGCGCTGTGCGGCGCACTTCG 1012
Db      781 GGTCCAGTCTCTGTCTCC--TCAAAGGACAGCAGAGCTTGTTCATGTTCCATTC-- 837
Oy      1013 ACATAAATACAGTATTCCTCACTCTTATCTTACAACTCCCGCAGCCCACTGCACTTC 1072
Db      838 -----ACAGACGTATCTTGTCTCTTAACTCCATCCACAACTATCCACCTTC 891
Oy      1073 ACTAGTCCCAATTCCTGACCTTGTGAGGCGGCGGAGTATCTGACCTCCCTCGGCA 1132
Db      892 ACTAGTCCCAATTCCTGACCTTGTGAGGCGGCGGAGTATCTGACCTCCCTCGGCA 936
Oy      1133 CAGACCCCGCAGGAGTGTCTTCACTGACTCTGTGGGCAAGATGGGTCCAGAAAGCC 1192
Db      937 CAGCCGACCACTGTTTATTTGACTTGTGTGAC----- 968
Oy      1193 CACTTACGACACTAAGAGGCGCTGAGCTGGCGGCGGAGAACCAAGAGACTGGCCCTAG 1252
Db      969 -----CAGGACTGAGATGGGCTGAGCTGTGGGAGGAGCAAGCTCGGACTGA 1023
Oy      1253 GCCAGAGTTCCTCAATTCGAGGAGGAGGAG--AACAGACAAAGCTCCCTTGAAGATT 1311
Db      1024 GCCAGAGTTCCTCACTGAGGAGGAGGAGGAGCTGGGAGCAAGCTCGGACTGA---TC 1079
Oy      1312 CCGTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1366
Db      1080 CCGTGTGATTTTGA--AGATACATTTTATTTATTTATTTATTTATTTATTTATTTA 1132

```

```

RESULT 10
LOCUS     AC126921
DEFINITION Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered
pieces.

```

```

ACCESSION AC126921
VERSION    AC126921.1
KEYWORDS   HTGS_PHASE1, HTGS_DRAFT.
SOURCE     HTGS, HTGS_PHASE1, HTGS_DRAFT.
ORGANISM   Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 15353)
REFERENCE  Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
AUTHORS   Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Latic, P.,
Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Mastello, C., Maskeri, B., Mastrian, S.D.,
McCluskey, J.C., McDowell, J., Pagurigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantirlop, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsurgren, C., Vogt, J.L., Walker, M.A.,
Wetherby, K.D., Wilgins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 15353)
Green, E.D.
Direct Submission
Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gailthersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoehghrl.nih.gov
----- Project Information
Center project name: ddi
Center clone name: 045D24
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 146066 bases at least Q40
Consensus quality: 147748 bases at least Q30
Consensus quality: 148824 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 152453; sum-of-coverage
Quality coverage: 8.80x in Q20 bases; agarose-fp
Quality coverage: 8.72x in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 2312: contig of 2312 bp in length
* 2313
* 2412: gap of unknown length
* 2413
* 5841: contig of 3429 bp in length
* 5842
* 5941: gap of unknown length
* 5942
* 8435: contig of 2494 bp in length
* 8436
* 8535: gap of unknown length
* 8536
* 15799: contig of 7264 bp in length
* 15800
* 15899: gap of unknown length
* 15900
* 25224: contig of 9325 bp in length
* 25225
* 32504: gap of unknown length
* 32525
* 32526: contig of 7180 bp in length
* 32505
* 32604: gap of unknown length
* 32605
* 40970: contig of 8366 bp in length
* 40971
* 41070: gap of unknown length
* 41071
* 56580: contig of 15520 bp in length
* 56591
* 73769: gap of unknown length
* 73769: contig of 17079 bp in length
* 73770
* 73869: gap of unknown length
* 90859: contig of 16990 bp in length

```

\* 90860 90959: gap of unknown length  
\* 90960 111428: contig of 20469 bp in length  
\* 111428 111528: gap of unknown length  
\* 111529 153553: contig of 42025 bp in length.  
Location/Qualifiers  
source 1. 153553

/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone="RP42-45D24"  
/clone\_11b="RP42"  
1. 2312  
misc\_feature /note="assembly\_fragment"  
2413. 5841  
/note="assembly\_fragment"  
clone\_end="r7  
vector\_side="left"  
5942. 8435  
misc\_feature /note="assembly\_fragment"  
8536. 15799  
/note="assembly\_fragment"  
15900. 25224  
/note="assembly\_fragment"  
23325. 32504  
/note="assembly\_fragment"  
32605. 40970  
/note="assembly\_fragment"  
41071. 56590  
/note="assembly\_fragment"  
56691. 73769  
/note="assembly\_fragment"  
73870. 90859  
/note="assembly\_fragment"  
90960. 111428  
/note="assembly\_fragment"  
111529. 153553  
/note="assembly\_fragment"  
clone\_end="SP6  
vector\_side="right"  
BASE COUNT 39178 a 37393 c 36616 g 39259 t 1107 others  
ORIGIN

Query Match 26.5%: Score 364.2; DB 2; Length 153553;  
Best Local Similarity 74.9%: Pred. No. 1.1e-56;  
Matches 588; Conservative 0; Mismatches 133; Indels 64; Gaps 8;  
QY 591 CCTGACGTGTCAGTGCCTTGATGAGGGGAAAGCTGTCTACCTGAAGCTGACTTGT 650  
DB 27179 CCTGACCTCCCGAGGTGACCTTTGATGAGGGGAAAGCTGTCTACCTGAAGCTGACTTGT 27238  
QY 651 GGTGATGATGTCGTGGCCCTGCTGCTGCTGAGGAAATTCAGCCACTCGGCCAGTTTC 710  
DB 27239 GGTGATGATGACAGCTGGCCCTGCTGCTGCTGAGGAAATTCAGCCACTCGGCCAGTTTC 27238  
QY 711 CCTGGGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGGCTC 770  
DB 27239 CCTGGGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTTGGCCCTGCGGCCAGGCTC 27358  
QY 771 CTCCCTGGGAGTCGACCTCCCTGGGCCCATCTCAAGGCTGCCCTCTCTACCTA 830  
DB 27359 TTCCCTGGGAGTCGACCTCCCTGGGCCCATCTCAAGGCTGCCCTCTCTACCTA 27418  
QY 831 CTTCGACTCTTCAGTTTCACTGAGGGCCCTTGCTCCCCACAGTGTGCCAGGCTC 890  
DB 27419 CTTCGACTCTTCAGTTTCACTGAGGGCCCTTGCTCC-----AGGTGG 27464  
QY 891 CGGCTCCCTCGACAGCTCTTGAGGCA-CCGCGTCCCTCTGCTCCGCCACCTCGAGCGCTC 949  
DB 27465 CTTCGCTCCAGCGCGCTCTGAGACACCTCTGCCCTGTGCTCCGCCACCTGAGCTGCTC 27524  
QY 950 TTTCGCTCCAGAGCTCCCTCTGAGAGGCTCGCTGGGCTGTTCACAGTGTTCCTC 1009  
DB 27525 CTTCGCTCCAGAGCTCCCTCTGAGAGGCTCGCTGGGCTGTTCACAGTGTTCCTC 27580

QY 1010 CCCACATAATACAGTATTTCCACCTCTTATCTTACAACTCCGCCACGCCACCTCCAC 1069  
DB 27581 GGCACACCAATATTTCCACCTC-----TTCCACCTCCACCTCCAC 27633  
QY 1070 CTCTAGCTCCCAATCCCTGACCTTTGAGGCCCCCAAGTATCTGACCTCCCTCTG 1129  
DB 27624 CTCACAGCTCCCGAGCCCATTTATCTCTGACCTCC-----CAACCTCG 27670  
QY 1130 CCACAGACCCCGAGGCACTTGCTTCACCTGCTACTCTGTGTGGGCAAGATGGCTCAGAA 1189  
DB 27671 CTGACAGCCCGAGGCACTGTTGACTGTACTCTG-GGTATATGGTCTCCATAC 27729  
QY 1190 CCCCACTTCAGGCACTAAGAGGGGCTGACCTGGCGGCGAGCAAGCAAGACTGGGCC 1249  
DB 27730 CCCCTCTTCAGGCACTAAGAGGGGCGGAT-----GCCAGAGACACAGGGAC 27776  
QY 1250 TAGGCCAGAGTTCCCAATGTGAGGGGCGAGA-AACAAGACAGCTCTCTCCCTTGAGA 1308  
DB 27777 TAGGCCAGAGTTCCCAATGTGAGGGGCGAGAAGACAGAGCAGAGCTCTCTCCCTGATPA 27836  
QY 1309 TTCCCTGCTGATTTTAAACAGATATTTATTTATTTATTTATTTGACAAATGTTGATA 1368  
DB 27837 TTCCCTGCTGATTTTAAACAGATATTTATTTATTTATTTATTTGAGACTTAAATGTTGATA 27896  
QY 1369 AATGG 1373  
DB 27897 ACTGG 27901

RESULT 11  
AC069459/c  
LOCUS  
DEFINITION Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT  
ACCESSION AC069459  
VERSION AC069459.23 GI:14547768  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.

REFERENCE  
AUTHORS  
Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 203083)  
Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,  
Dederich, D., Thomas, S., Okunou, G., Carlod, C., Garner, T.,  
Addison, S., Pace, A., Williams, G., Bonini, D., Brooks, A., Brown, J.,  
Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,  
Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Koch, S.,  
Fernandez, C., Ferraguto, D., Forcum-Tansey, D., Gill, R.,  
Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogue, M.,  
Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,  
Kovar, C., Liu, J., Liu, W., Louisse, H., Lozano, R.J., Martin, R.,  
Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,  
Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogun, M., Farish, B.,  
Perez, L., Reller, D., Say, J., Shen, H., Vasquez, L., Watling, S.,  
Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A.,  
Munzy, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein, G.,  
Worley, K. and Glbbs, R.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 203083)  
AUTHORS Worley, K.C.  
TITLE Direct Submission  
JOURNAL Direct Submission  
COMMENT Submitted (31-May-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jun 25, 2001 this sequence version replaced gi:12621364.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: MAFO





Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 112720 bases at least Q40  
Consensus quality: 116666 bases at least Q30  
Consensus quality: 119165 bases at least Q20

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.bhsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.bhsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1022: contig of 1022 bp in length  
1023 1122: gap of unknown length  
1123 2394: contig of 1272 bp in length  
2395 2494: gap of unknown length  
2495 3663: contig of 1169 bp in length  
3664 3763: gap of unknown length  
3764 4879: contig of 1116 bp in length  
4880 4979: gap of unknown length  
4980 6787: contig of 1808 bp in length  
6788 8192: contig of 1305 bp in length  
8193 8293: gap of unknown length  
8293 9461: contig of 1169 bp in length  
9462 9561: gap of unknown length  
9562 11435: contig of 1874 bp in length  
11436 11535: gap of unknown length  
11536 13535: contig of 2000 bp in length  
13536 17323: gap of unknown length  
17324 17423: gap of unknown length  
17424 20406: contig of 2983 bp in length  
20407 20506: gap of unknown length  
20507 23445: contig of 2939 bp in length  
23446 23545: gap of unknown length  
23546 25600: contig of 2055 bp in length  
25601 25701: gap of unknown length  
25702 28049: contig of 2349 bp in length  
28050 28149: gap of unknown length  
28150 30474: contig of 2325 bp in length  
30475 30574: gap of unknown length  
30575 35498: contig of 4924 bp in length  
35499 35598: gap of unknown length  
35599 39284: contig of 3686 bp in length  
39285 39384: gap of unknown length  
39385 41970: contig of 2586 bp in length  
41971 42070: gap of unknown length  
42071 45659: contig of 3589 bp in length  
45660 45759: gap of unknown length  
45760 50332: contig of 4573 bp in length  
50333 50432: gap of unknown length  
50433 54695: contig of 4263 bp in length  
54696 54795: gap of unknown length  
54796 58289: contig of 3494 bp in length  
58290 58388: gap of unknown length  
58389 63599: contig of 5210 bp in length  
63600 70305: contig of 6606 bp in length  
70306 70405: gap of unknown length  
70406 76123: contig of 5718 bp in length  
76124 76223: gap of unknown length  
76224 84961: contig of 8738 bp in length  
84962 85061: gap of unknown length  
85062 93614: contig of 8553 bp in length  
93615 93714: gap of unknown length  
93715 103352: contig of 9638 bp in length  
103353 103452: gap of unknown length

FEATURES  
source 1. 138792  
/organism="Rattus norvegicus"  
/db\_xref="taxon:10116"  
/clone="CH230-320N23"

BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others

ORIGIN

Query Match 18.5% Score 254.2; DB 2; Length 138792;  
Best Local Similarity 66.7%; Pred. No. 1.2e-36;  
Matches 517; Conservative 0; Mismatches 188; Indels 70; Gaps 8;

QY 601 CAGGTGCACTTGTATGAGGGGAAAGGCTCTACCTGAAGCTGACCTGTGTGATGCT 660  
|||||  
Db 137627 CAGGTGCACTTGTATGAGGGGAAAGGCTCTACCTGAAGCTGCTGTGTGATGCT 137686  
|||||

QY 661 GTCTGCGCCCTGCGCTGCTGAGAGAAATCTCAGCCACTGCGGCAATTCCTCGGCCC 720  
|||||  
Db 137687 GTCTGCGCCCTGCGCTGCTGAGAGAAATCTCAGCCACTGCGGCAATTCCTCGGCCC 137746  
|||||

QY 721 CAGCTCGCCCTGCTGAGAGAGTCTGAGGCTGTGAGCCCTGCGGCAAGGCTCTCGTCCG 780  
|||||  
Db 137747 CAGCTCGCCCTGCTGAGAGAGTCTGAGGCTGTGAGCCCTGCGGCAAGGCTCTCGTCCG 137806  
|||||

QY 781 ATCCGCAACCTCCCTGAGGCCATCTCAAGGCTGCCCCCTTCTCAGCTTCTGCACTC 840  
|||||  
Db 137807 ATCCGCAACCTCCCTGAGGCCATCTCAAGGCTGCCCCCTTCTCAGCTTCTGCACTC 137866  
|||||

QY 841 TTCGAGGTACAGTAGGGGCGCTGCT-----CTCCCAAGTGTCTCCAGGCTGCG 892  
|||||

Db 137867 CACGAAAGCAAGAGAGATCTGATATGACAGACTTCTTATCTCCAACTAAGATA 137926  
|||||

QY 893 GCTCCCTGCAAGCTCTGAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952  
|||||  
Db 137927 CCAAGACCAAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137986  
|||||

QY 953 GCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012  
|||||  
Db 137987 GCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138045  
|||||

QY 1013 ACATAAATACAGTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072  
|||||  
Db 138046 ACAGACA-----TATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138100  
|||||

QY 1073 ACTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132  
|||||  
Db 138101 ACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138138  
|||||

QY 1133 CAGACCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192  
|||||  
Db 138139 CCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138171  
|||||

QY 1193 CACTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252  
|||||  
Db 138172 TGCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138231  
|||||

QY 1253 GCCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1311  
|||||  
Db 138232 GCCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138287  
|||||

QY 1312 CCTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTT 1366  
|||||  
Db 138288 CCTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTT 138340  
|||||

RESULT 14

AC130192/c	AC130192	177555 bp	DNA	linear	HTG 08-AUG-2002							
DEFINITION	Sus scrofa clone RP44-436K21, WORKING DRAFT SEQUENCE, 10 unordered pieces.											
ACCESSION	AC130192											
VERSION	AC130192.1	GI:22138443										
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.											
SOURCE	pig.											
ORGANISM	Sus scrofa											
REFERENCE	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.											
AUTHORS	1 (bases 1 to 177555) Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hashighi, P., Hansen, N., Ho, S.-L., Iodl, J.R., Karlins, E., Laric, P., Lee-Jin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Magallanes, E.H., Mastello, C., Masker, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pagulirigan, C., Pearson, R., Potinoy, M.E., Prasad, A., Scheller, M.G., Stantipour, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgouen, C., Vogt, J.L., Walker, M.A., Welberly, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.											
TITLE	NISC Comparative Sequencing Initiative											
JOURNAL	Unpublished											
REFERENCE	2 (bases 1 to 177555)											
AUTHORS	Green, E.D.											
TITLE	Direct Submission											
JOURNAL	Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA											
COMMENT	----- Genome Center											

```

Center clone name: 436K21
-----
Summary statistics
Sequencing vector: plasmid; N/A; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 164120 bases at least Q40
Consensus quality: 168085 bases at least Q30
Consensus quality: 170767 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 176655; sum-of-contigs
Quality coverage: 4.78x in Q20 bases; agarose-fp
Quality coverage: 4.57x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 10 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1
3460: contig of 3460 bp in length
3461
3560: gap of unknown length
3561
6869: contig of 3309 bp in length
6870
6970: gap of unknown length
6970
6970
13466: contig of 6496 bp in length
13466
13565: gap of unknown length
13566
21301: contig of 7736 bp in length
21302
21401: gap of unknown length
21402
33719: contig of 12318 bp in length
33720
33819: gap of unknown length
33820
45054: contig of 11255 bp in length
45055
45155: gap of unknown length
45155
58724: contig of 13570 bp in length
58724
58824: gap of unknown length
58825
76670: contig of 17846 bp in length
76671
76670: gap of unknown length

```

FEATURES	
*	76771 115165: contig of 38395 bp in length
*	115166 115265: gap of unknown length
*	115266 177555: contig of 62290 bp in length.
	Location/Qualifiers

```
misc_feature 1..3460
note="assembly-fragment
3561..6869
/note="assembly-fragment
6970..13465
/note="assembly-fragment"
13566..21301
/note="assembly-fragment"
21402..33719
/note="assembly-fragment"
33820..45054
/note="assembly-fragment"
45155..58724
/note="assembly-fragment
clone_end:77
vector_side:left"
58825..76670
/note="assembly-fragment
76771..115165
/note="assembly-fragment"
115266..177555
```

	vector_side:right"
BASE COUNT	41026 a 46997 c 46681 g 41924 t 927 others
ORIGIN	

Query Match	17.9%	Score 246.4;	DB 2;	Length 177555;
Best Local Similarity	85.6%	Pred. NO. 3.2e-35;		
Matches 274;	Conservative 0;	Mismatches 46;	Indels 0;	Gaps 0;

[illegible]

RESULT	15
AC126925/c	
LOCUS	161428 bp DNA linear HTG 10-JUL-2002
DEFINITION	Canis familiaris clone Rpa1-332E11, WORKING DRAFT SEQUENCE, 15 unordered pieces.
ACCESSION	AC126925
VERSION	AC126925.1 GI:21724102
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4-P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:59 ; Search time 42.8782 Seconds  
(without alignments)  
1081.216 Million cell updates/sec

Title: US-09-245-198a-2  
Perfect score: 1162  
Sequence: 1 VLISGLALACIGLLVVVSL.....PMAHLKAPFLTYGLRQVH 225

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	107.5	9.3	325	5 Q9V5G2	Q9V5G2 drosophila
2	93.5	8.0	557	16 Q8XOX3	Q8XOX3 ralsstonia s
3	90.5	7.8	210	16 Q9A926	Q9A926 caulobacter
4	90.5	7.8	421	16 Q9HWM2	Q9HWM2 pseudomonas
5	89.5	7.7	805	5 Q9VPD4	Q9VPD4 drosophila
6	89	7.7	287	13 Q9WMP9	Q9WMP9 gallus gall
7	89	7.7	352	12 Q89341	Q89341 hendra viru
8	88	7.6	224	5 Q9V762	Q9V762 drosophila
9	88	7.6	353	12 Q66760	Q66760 equine morh
10	88	7.6	532	4 Q16727	Q16727 homo sapien
11	88	7.6	1663	4 Q9U001	Q9U001 homo sapien
12	88	7.6	3261	4 Q9V556	Q9V556 homo sapien
13	88	7.6	3664	4 Q96T58	Q96T58 homo sapien
14	87.5	7.5	522	10 Q9FTN7	Q9FTN7 oryza sativ
15	87.5	7.5	670	16 Q9AA15	Q9AA15 caulobacter
16	86.5	7.4	340	16 Q9HUR8	Q9HUR8 pseudomonas

17	86.5	7.4	504	16 Q92KA4	Q92KA4 rhizobium m
18	86	7.4	260	10 Q8S2N9	Q8S2N9 oryza sativ
19	85.5	7.4	331	10 Q942P9	Q942P9 oryza sativ
20	84.5	7.3	422	16 Q9RKB0	Q9RKB0 streptomyce
21	84.5	7.3	430	2 Q9RE01	Q9RE01 streptomyce
22	84.5	7.3	1079	13 Q8UVR4	Q8UVR4 xenopus lae
23	84.5	7.3	1118	16 Q98E34	Q98E34 rhizobium l
24	84	7.2	2962	5 Q93326	Q93326 caenorhabdi
25	83.5	7.2	394	16 Q92V66	Q92V66 rhizobium m
26	83	7.1	467	16 Q9S2Y4	Q9S2Y4 streptomyce
27	83	7.1	2779	5 Q9W4N7	Q9W4N7 drosophila
28	82.5	7.1	174	16 Q9CKX1	Q9CKX1 pasteurella
29	82.5	7.1	314	5 Q8WPH7	Q8WPH7 theileria a
30	82.5	7.1	549	16 Q8RC38	Q8RC38 thermoaer
31	82	7.1	619	5 Q8SU9	Q8SU9 encephalito
32	82	7.1	654	16 P74664	P74664 synechocyst
33	81.5	7.0	999	11 Q9JUR6	Q9JUR6 mus musculu
34	81.5	7.0	1665	11 Q8V111	Q8V111 mus musculu
35	81	7.0	258	4 Q9NDD0	Q9NDD0 homo sapien
36	81	7.0	262	11 Q9CVL1	Q9CVL1 mus musculu
37	80.5	6.9	324	11 Q9JUS4	Q9JUS4 beta vulgar
38	80.5	6.9	788	4 Q9ZES5	Q9ZES5 rattus norv
39	80.5	6.9	1438	16 Q8S739	Q8S739 pseudomonas
40	80.5	6.9	319	16 Q8U8B6	Q8U8B6 agrobacteri
41	80	6.9	403	6 Q9BGL8	Q9BGL8 canis famli
42	80	6.9	589	5 Q8T0B2	Q8T0B2 drosophila
43	80	6.9	793	5 Q9Y108	Q9Y108 drosophila
44	80	6.9	793	5 Q9Y108	Q9Y108 oryza sativ
45	80	6.9	1522	10 Q8S4C3	Q8S4C3 oryza sativ

## ALIGNMENTS

### RESULT 1

Q9V5G2	PRELIMINARY;	PRT;	325 AA.
ID Q9V5G2			
AC Q9V5G2:			
DT 01-MAY-2000 (TREMBLrel. 13, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE CG12919 protein.			
GN CG12919.			
OS Drosophila melanogaster (Fruit fly).			
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC Ephydroidea; Drosophilidae; Drosophila.			
OX NCBI_TaxID=7227;			
RA [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=BERKELEY;			
RX MEDLINE=20196006; PubMed=10731132;			
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,			
RA Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,			
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Moperson D.,  
 RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Piltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spadling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gidbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AEO03831; AAF58648.1; -  
 DR FlyBase: FBgn0033483; CG12919.  
 DR InterPro: IPR000478; TNF\_family.  
 DR SMART: SM00207; TNF\_1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 SO SEQUENCE 325 AA; 35862 MW; 6E5CCB6964F1A3A CRC64;

Query Match 9.3%; Score 107.5; DB 5; Length 325;  
 Best Local Similarity 23.1%; Pred. No. 0.046;  
 Matches 48; Conservative 34; Mismatches 91; Indels 35; Gaps 9;

QY 29 QEPSEELTADREPELNPQTEESQDVVPLEQLVPRRS--APGRKAPRAIAAH 86  
 DB 141 QEKSNATKESPPAPLHRRMRSH-----RHLYRKGESLTSARSEDSRP---AAH 191  
 QY 87 YEVPREPQDGAQGVDTGSGMEETKINSSPLRYDQGEFTVAGLYYLCOVHP 146  
 DB 192 FHLSRRRHGSM-GYHGDMYIGNDNERNISYQ--HFQTRGVLTVTGTLVYVAQICYN 249  
 QY 147 EGRAVYTKLDLVNGVLA-----LRCLFEFSATTAASPEPQRLQCVSGLLPLRGSS 199  
 DB 250 NSHD-----QNGFTVFQGDTPFLQCLN---TVPTNPMHKVHTCHTSGLLHLERNER 297  
 QY 200 LRITL---PWAKKAPFLTYFGLFOV 224  
 DB 298 IHLKDINHNRNAVLRGNNSYFGLFV 325

RESULT 2  
 OSXOX3 PRELIMINARY; PRT; 557 AA.

AC OSXOX3;  
 DT 01-MAR-2002 (TREMblrel. 20, Created)  
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
 DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Probable activation/secretion signal peptide protein.  
 GN RSP1095 OR RS02601.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OC NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GM1100;  
 RX MEDLINE-21681879; Pubmed-11823852;  
 RA Sclanoubet M., Genin S., Artiguenave F., Gouzy J., Manenot S.,  
 RA Arlet M., Billault A., Brotter P., Camus J.C., Catolico L.,  
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weisenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646082; CAD18246.1; -.

KW Plasmid; Complete proteome.  
 SO SEQUENCE 557 AA; 60786 MW; 8B974C147D710649 CRC64;

Query Match 8.0%; Score 93.5; DB 16; Length 557;  
 Best Local Similarity 26.8%; Pred. No. 2;  
 Matches 62; Conservative 25; Mismatches 77; Indels 69; Gaps 14;

QY 2 LSLGLALACGLLVVSLGSMATLS-AQEPSEELTADREPE--LNPQTEESQDVV 58  
 DB 10 LALGVALA-----GFLSSMAQSSPQGNPLDTLPVDTSPRPCKIHVQGRN--- 58  
 QY 59 PTEQL-----VPRRSAPKGRKAPRAIAAHYVHPREPQDGAQGVDTGSGMEETKI 114  
 DB 59 PALENLTATLTPESKQIEGVKALPPEEIAHFA--PMAGHDVTVAQLQAA--EVTXL 114  
 QY 115 --NASSPLRYDQGEFTVAGLYYLCOVHDEKAVYTKLDLVNGVLAIRCLEERS 172  
 DB 115 YADRGYPL-----SFAFPAQTF-----EGGVARI---TVEGVIAKMRIE--- 152  
 QY 173 ATAASSPFP---QLRLC-----QVSGLLPLRGSSLRITLP 206  
 DB 153 ----GKRGPLEARLRATSKHMDERPLRRTFERYVTGVALQGVQITATVP 201

RESULT 3  
 Q9A926 PRELIMINARY; PRT; 210 AA.  
 AC Q9A926;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Hypothetical protein CC1168.  
 GN CC1168.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OC NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; Pubmed-11259647;  
 RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadok N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouir H., Shetty J., Berry K.,  
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005794; AAK23152.1; -.  
 DR TIGR: CC1168; -.  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 210 AA; 22344 MW; 8B9830ADF87F45C CRC64;

Query Match 7.8%; Score 90.5; DB 16; Length 210;  
 Best Local Similarity 26.2%; Pred. No. 1.1;  
 Matches 33; Conservative 19; Mismatches 67; Indels 7; Gaps 2;

QY 2 LSLGLALACGLLVVSLGSMATLSAQEPSEELTADREPELNPQTEESQDVVPL 61  
 DB 32 LRLGLAVFAGVATVIVQAMRSCLASSKLOTAAAPVLDKPRP---TGVLKDRPFL 88  
 QY 62 EQLVPRRSAPKGRKAPRAIAAHYVHPREPQDGAQGVDTGSGMEETKINS 117  
 DB 89 ITTERARERDAKDONIVRLTAPLVKGYGEPNPSQATKASGVYREAEVTLTLDVEVATTS 148  
 QY 118 SPLRYD 123  
 DB 149 EGDFD 154

RESULT 4

Q9H0W2 PRELIMINARY; PRT: 421 AA.  
 AC Q9H0W2;  
 DT 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)  
 DE Hypothetical protein PA4851.  
 GN PA4851.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas  
 NCBI\_TaxID=287;  
 RN NCBI\_TaxID=287;  
 RP [1]  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Gardner R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lam R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sliet M.H., Hancock R.E.W., Lory S., Olson M.V.,  
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen."  
 RL Nature 406:959-964(2000).  
 DR EMBL; AE004898; AAC08236.1;  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 421 AA; 46442 MW; 68C5C17099953C3E CRC64;

Query Match 7.8%; Score 90.5; DB 16; Length 421;  
 Best Local Similarity 22.6%; Pred. No. 2.7;  
 Matches 53; Conservative 27; Mismatches 76; Indels 79; Gaps 10;

OY 29 QEPSEELTADR-----REPELNPQTESQDVVPL-----EQLV 65  
 DB 163 QEPSEGLPLAERVSFVPLRAEDPELRLSDDDDEPSGKRRHGLSALDEDESSA 222  
 OY 66 RPRRSAPGRKARPRRAIAAHYVHPRPGODGAGVGTSGWETKINSSPLRYDRQ 125  
 DB 223 EEGSVAPLGSARKPDEA-----PVESLDQLRDEPLQLAWMKPR-----RQWRPR 266  
 OY 126 IGEFTVI---RAGLYYYLYCQVHFE-----GKAVYLIKLDL--LVNGV 162  
 DB 267 LGWLLILLILALGGLAAQYIAVHFDELARQDARVPMFAOLCEPICGCTLSKVDVEQIRSSN 326  
 OY 163 LALRCLTEESA-----TASSPGP--QLRLCQVSGLL-----PLRPGSSL 200  
 DB 327 LVKSHPEFSGALVVDATITNRASFSPPLLELRFADLNGHLJANRRFPGEYL 381

RESULT 5  
 O9VFD4 PRELIMINARY; PRT: 805 AA.  
 AC O9VFD4;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE CG18442 protein.  
 GN CG18442.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek J.R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Abganyant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup I.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Klip D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,  
 RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003707; AA055126.1;  
 DR FlyBase; FBgn0038287; CG18442.  
 SQ SEQUENCE 805 AA; 86321 MW; FDF9FEFF14E69957 CRC64;

Query Match 7.7%; Score 89.5; DB 5; Length 805;  
 Best Local Similarity 21.7%; Pred. No. 7.7;  
 Matches 49; Conservative 27; Mismatches 55; Indels 95; Gaps 10;

OY 29 QEPSEELTADRREPELNPQTESQDVVPLLEQVHPRRSAPGRKARPR--ATAAH 86  
 DB 482 EQPGQEEI-----QPPRPTPTTE-----LSPPARPPSAELLQRYSPKKQVRIAS 529  
 OY 87 YEYHRR-----PGODGAGVDTGYSGEHKKI-----N 115  
 DB 530 PYMOPQERRELCPPLPPRGSPITLDGSSSPTNAVSGKPKPLPIACRPRPSNGVSPN 589  
 OY 116 SSS-----PLRYDRI-----GFTVIRAGLYLYCOV 143  
 DB 590 SSSPGSAPPANTSPPIPTVRLPHLNQANGTLPLLPKPPQOLHGKFLIKNG----- 641  
 OY 144 HFDEGKAVYLIKLDLVLNGVL-----ALRCLSEFSATAASPGPOL 183  
 DB 642 HLIDDEALLAKTDVAMSGLLIKLDQVAQC-----SVAQAAGGRTSI 683

RESULT 6  
 O90WT9 PRELIMINARY; PRT: 287 AA.  
 AC O90WT9;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE TNF-related apoptosis inducing ligand.  
 OS Gallus gallus (Chicken).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 CC Gallus  
 NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA BrIdgham J.T., Johnson A.L.;  
RT "TNF-related apoptosis inducing ligand (TRAIL) expression in the hen  
ovary.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY057941; AAL23702.1; -  
DR InterPro: IPR003636; TNF\_abc.  
DR pfam: PFO0229; TNF; 1.  
DR ProDom: PDD00212; TNF\_abc; 1.  
DR PROSITE: PS50049; TNF-2; 1.  
SQ SEQUENCE 287 AA; 32092 MW; DB06ELC95087B108 CRC64;

Query Match 7.7%; Score 89; DB 13; Length 287;  
Best Local Similarity 23.4%; Pred. No. 2.3;  
Matches 40; Conservative 24; Mismatches 65; Indels 42; Gaps 7;

QY 84 AAHYEVHPR-PGQDGAAGADG-----TVSGWEERKINS-SPLRYDQIGEPYIRA 134  
DB 127 SAHLIRPQNPADGSSRRFGNLSQCRHAITWEDSTHSHLQNTY--RDGRLRVNQA 184  
QY 135 GLLYLQVHF-----EGKAVYLRKLDLVNGVLAIRLCEFSATA 175  
DB 185 GKYYVSQIFRYSRQAGARVSPOLVOCINMKTYSQPIILLKCV-----GTK 234  
QY 176 ASGPGQ--LRLCQVGLRPRGSSLRIRTPMAHLKAAPLTYGGLFOV 224  
DB 235 CMAPEAEYGLHALYOGGLFELKAGDELIVSVSLAIDSPDAASYGAFRL 285

## RESULT 7

ID 089341 PRELIMINARY; PRT; 352 AA.  
AC 089341;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Matrix.  
GN M.  
OS Hendra virus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
OX NCBI\_TaxID=63330;  
RN RN  
RP SEQUENCE FROM N.A.  
RA MEDLINE=90581172;  
RA Yu M., Hansson E., Langedijk J.P., Eaton B.T., Wang L.F.;  
RT "The attachment protein of Hendra virus has high structural similarity  
but limited primary sequence homology compared with viruses in the  
RT genus Paramyxovirus.";  
RL Virology 251:227-233(1998).  
RN RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20481636; PubMed=11024125;  
RA Wang L.F., Yu M., Hansson E., Pritchard L.I., Shnell B.,  
RA Michalski W.P., Eaton B.T.;  
RT "The exceptionally large genome of Hendra virus: support for creation  
of a new genus within the family Paramyxoviridae.";  
RL J. Virol. 74:9972-9979(2000).  
RN RN  
RP SEQUENCE FROM N.A.  
RA Wang L.F., Yu M., Pritchard L.I., Hansson E., Eaton B.T.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF017149; AAC83191.1; -  
DR InterPro: IPR000982; Matrix.  
DR pfam: PFO0661; Matrix; 1.  
DR ProDom: PDD00741; Matrix; 1.  
SQ SEQUENCE 352 AA; 39793 MW; 79E238DE496828D5 CRC64;

Query Match 7.7%; Score 89; DB 12; Length 352;  
Best Local Similarity 28.1%; Pred. No. 3;  
Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;

QY 78 RPRRAIAHYEHPRGQDGAQVDTVSGWEETKINSSPLRYDQIGETVIRAGLY 137

DB 197 RRNNIAIFNLVLYLKIDADLAKIGSFDK-DGTVAS-----FMLHIGNF-VRRAGXY 249  
QY 138 YL--YQVHFDECKAY-----LKLIDLNVNGVLAIRLCEFSATRAASSPGQLRQV 188  
DB 250 YSVEYCKRKIDRKLQFSGISGGLSLTKINKVSKRULPAQM-----GLQKMLC-- 299  
QY 189 SGLPLRPGSSLRIRTPWP 207  
DB 300 FSLMDINPM---LNRLLTW 314

## RESULT 8

ID 09V762 PRELIMINARY; PRT; 224 AA.  
AC 09V762;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE CG10205 protein.  
GN CG10205.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agdayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusslein-Kolbe J., Pacle J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein D.A., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003813; AAF58200.1; -  
DR Flybase: FBgn003970; CG10205.  
SQ SEQUENCE 224 AA; 25025 MW; 780B9EF7BBAFF156 CRC64;

Query Match 7.6%; Score 88; DB 5; Length 224;



```

OY 74 ---GRKARPRRAIAAHYEVHPRPGDGAAGVGTSGMEETKINSSPLRYDROIGFT 130
DB 1459 PTTSGSTPGGLVLPHTTEFPAPKQDSS-----PILTSQRPVDMVQLLKKYP 1505
OY 131 VIRAGLYYL-----YCOVHFDEGKAVYLLKLDLVNGVLAIRCLFEESATTAASPGQLRL 185
DB 1506 IYWQGLLALKNDTAAYQLHVFVSG-----NNVLAHRSLL-----PLSGGPPRLRI 1548
OY 186 CQ 187
DB 1549 AQ 1550

RESULT 12
OY9556 PRELIMINARY: PRT: 3261 AA.
AC 09Y556:
DB 01-NOV-1999 (TREMBLrel. 12, Created)
DB 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 357.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCB1_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Rhodes S., Huckle E.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL096858; CAB51072.1;
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; rrm; 2.
DR PROSITE: PS50102; RRM; 2.
KW Hypothetical protein.
SQ SEQUENCE 3261 AA: 357037 MW: F8B6A645DD9B6BC CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3261;
Best Local Similarity 23.1%; Pred. No. 65;
Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

OY 33 QEELTADREPELNPQTEESQ-----DVYPLEQLVRRRSAPK----- 73
DB 2998 QPRLPGPANRPEPHPTQVQRAQETGPTSPSPVSVSMKPDLPVSLPTQTAPKQPLEV 3056
OY 74 ---GRKARPRRAIAAHYEVHPRPGDGAAGVGTSGMEETKINSSPLRYDROIGFT 130
DB 3057 PTTSGSTPGGLVLPHTTEFPAPKQDSS-----PILTSQRPVDMVQLLKKYP 3103
OY 131 VIRAGLYYL-----YCOVHFDEGKAVYLLKLDLVNGVLAIRCLFEESATTAASPGQLRL 185
DB 3104 IYWQGLLALKNDTAAYQLHVFVSG-----NNVLAHRSLL-----PLSGGPPRLRI 3146
OY 186 CQ 187
DB 3147 AQ 3148

RESULT 13
OY9558 PRELIMINARY: PRT: 3664 AA.
AC 09Y558:
DB 01-DEC-2001 (TREMBLrel. 19, Created)
DB 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nuclear receptor transcription cofactor.
GN SHARE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCB1_TaxID=9606;
RN 11

```

```

RP SEQUENCE FROM N.A.
RA Shi Y., Downs M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,
RA Hon M., Evans R.M.;
RT "SHARP, an inducible cofactor that integrates nuclear receptor
RT repression and activation."
RL Genes Dev. 0:0-0(2001).
DR EMBL; AF356524; AAK52750.1;
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 4.
DR PROSITE: PS50102; RRM; 4.
KW Receptor.
SQ SEQUENCE 3664 AA: 402243 MW: 5228C58533E5B27B CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3664;
Best Local Similarity 23.1%; Pred. No. 75;
Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

OY 33 QEELTADREPELNPQTEESQ-----DVYPLEQLVRRRSAPK----- 73
DB 3401 QPRLPGPANRPEPHPTQVQRAQETGPTSPSPVSVSMKPDLPVSLPTQTAPKQPLEV 3459
OY 74 ---GRKARPRRAIAAHYEVHPRPGDGAAGVGTSGMEETKINSSPLRYDROIGFT 130
DB 3460 PTTSGSTPGGLVLPHTTEFPAPKQDSS-----PILTSQRPVDMVQLLKKYP 3506
OY 131 VIRAGLYYL-----YCOVHFDEGKAVYLLKLDLVNGVLAIRCLFEESATTAASPGQLRL 185
DB 3507 IYWQGLLALKNDTAAYQLHVFVSG-----NNVLAHRSLL-----PLSGGPPRLRI 3549
OY 186 CQ 187
DB 3550 AQ 3551

RESULT 14
OY9577 PRELIMINARY: PRT: 522 AA.
AC 09Y577:
DB 01-MAR-2001 (TREMBLrel. 16, Created)
DB 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative Rer1a protein (AtRER1A).
GN P0005A05.12.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriaridaceae; Oryzaceae; Oryza.
ON NCB1_TaxID=4530;
RN 11
RP SEQUENCE FROM N.A.
RA STRAIN-CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0005A05."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002863; BAB16908.1;
DR InterPro: IPR004932; Rer1.
DR Pfam: PF03248; Rer1; 1.
SQ SEQUENCE 522 AA: 58171 MW: AE43E294DFE8BE7F CRC64;

Query Match 7.5%; Score 87.5; DB 10; Length 522;
Best Local Similarity 30.3%; Pred. No. 6.9;
Matches 57; Conservative 21; Mismatches 57; Indels 53; Gaps 15;

OY 70 SAPKSGKARPRRAIAAHYEVHPRPGDGAAGVGTSGMEETKINSSPLRYDROIGFT 113
DB 201 SSSKGRGEFARVVA---SPSGFRPGFSMDAASGTAGAAAATAAKMTDASRAFY 256
OY 114 INSSPLRYDROIGFTVIRAGLYYLCOVHFDEGKAVYLLKLDLVNGVLAIRCLFEESATTAASPGQLRL 166
DB 257 LDRSTHATGRWIGTILAV--AAITAL--RYVLQGGYIYTYGGITL--LNLTL--GFLSPM 310
OY 167 CLEEFSAATAASPGQLRLCOVSGLLPLRPGSSLR--IRTLR-----WAHLKA---APPL 216

```



DB 311 VDPEAH-AASSDGP-----ALPTRGSDPEKPFIRLRPEPKRWAITKAFILAFVM 360  
 QY 217 TYFGLFQV 224  
 DB 361 TFSVFVDV 368

## RESULT 15

Q9AA15 PRELIMINARY; PRT; 670 AA.  
 AC Q9AA15;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)  
 DE Alpha-glucosidase, putative.  
 GN CC0796.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hatt D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Salterberg S.L., Venter J.C., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA "Complete genome sequence of Caulobacter crescentus."  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005755; AK22781.1; .  
 DR TIGR; CC0796; .  
 KW Complete proteome.  
 SQ SEQUENCE 670 AA; 72724 MW; CBEC8FF95B9300C2 CRC64;

Query Match 7.5%; Score 87.5; DB 16; Length 670;

Best Local Similarity 24.0%; Pred. No. 9.5; Mismatches 43; Conservative 24; Indels 37; Gaps 6;

QY 58 VPFLDQLVNP-----RRSAPKGRKR-----PRRAIAHYEVHPR----- 93  
 DB 498 LPFTRLGPMQYTPGGRNVAPKDEKSGFILPTVQTRGQALAMVYVDSPLTNWSDSP 557  
 QY 94 -GDDGAGQACVD--GTVSGMEETKINSSPLRYDRQIGETVI--RAGLYLYXCOVHDE 147  
 DB 558 IYVAASPACLDIFISAVPTSMDETRVLSG-----EIGQYIYARRKGADWWVGAMTNEA 610  
 QY 148 GKAVYLIKLDLNVGYLALCLEEFSATASPGPOLRCOVSGLLPLRPGSSLRIRTLTP 206  
 DB 611 GRTIVKVPFLSLNGAFSAEIRREGAEPYALKTRTQSVASKDTLTLLKLPAGGGGVIRISP 669

Search completed: March 31, 2003, 07:46:15  
 Job time : 56.8782 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 01:56:48 ; Search time 1488.36 seconds

(without alignments)  
2448.323 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSGLALACGLLVVSL.....PMAHLKAPFLYFGFQVH 225

Scoring table:

BLOSUM62		
Xgapop 10.0 , Xgapext 0.5		
Ygapop 10.0 , Ygapext 0.5		
Fgapop 6.0 , Fgapext 7.0		
Delop 6.0 , Delext 7.0		

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlh  
-O=/qgn2.1/USPFO/US09245198/runat\_24032003.163554.25956/app.query.fasta\_1.846  
-DB=EST -QFMT=fastap -SUFITX=first -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blissum62 -FRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09245198 @CGN.1.1.2463 @runat.24032003.163554.25956 -NCPU=6 -ICPU=3  
-NO\_XIPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST :	
1: em_estba :	
2: em_esthum :	
3: em_estlin :	
4: em_estlmu :	
5: em_estlov :	
6: em_estlpl :	
7: em_estro :	
8: em_hlc :	
9: gb_est1 :	
10: gb_est2 :	
11: gb_hlc :	
12: gb_est3 :	
13: gb_est4 :	
14: gb_est5 :	
15: em_estfun :	
16: em_estom :	
17: gb_gss :	
18: em_gss_hum :	
19: em_gss_inv :	
20: em_gss_pln :	
21: em_gss_vrt :	
22: em_gss_fun :	
23: em_gss_mam :	
24: em_gss_mus :	
25: em_gss_other :	
26: em_gss_pro :	
27: em_gss_rtd :	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1039	89.4	918	12	BF577781 602092080
2	968.5	83.3	731	13	B1871711 603395825
3	941.5	81.0	728	13	B1870393 603395641
4	882	75.9	777	13	B1819200 603034614
5	861	74.1	828	13	B1596681 603243254
6	861	74.1	834	13	B1766766 603056866
7	857	72.8	561	10	AM763237 ut70d09.y
8	831.5	71.6	948	14	B0707185 AGENCOURT
9	818	70.4	940	14	B0884231 AGENCOURT
10	764	65.7	963	14	B0671259 AGENCOURT
11	754	64.9	1033	11	AK020909 Mus muscu
12	721	62.0	1071	14	BM921213 AGENCOURT
13	703	60.5	951	14	B0674188 AGENCOURT
14	699	60.2	584	10	AM917574 EST348878
15	670	57.7	445	9	AA870722 vq25907.r
16	632.5	54.4	471	9	AA221610 my18d09.r
17	629	54.1	650	12	BG404836 602420160
18	621	53.4	360	10	BE654876 UI-M-BHO-
19	610	52.5	785	13	B1762908 BF041509 BP250025B
20	561.5	48.3	542	12	BF044430 BP250013A
21	549.5	47.3	404	12	BF466521 UI-M-CGP
22	541	46.6	315	12	BF466521 UI-M-CGP
23	513	44.1	474	13	B1965174 1d34e07.y
24	506.5	43.6	493	13	BE307031 601087888
25	501.5	43.2	894	13	BI908274 603068526
26	486	41.8	558	13	BM484863 538865 MA
27	477	41.0	581	13	B1738634 603358846
28	473	40.7	538	12	BF821434 MRL-RT003
29	458	39.4	298	14	BM688946 UI-E-C01-
30	429	36.9	278	10	AA417023 52923 MAR
31	428	36.8	342	9	AA637970 vrc30a10.r
32	404.5	34.8	571	12	BF073881 BF073881
33	404	34.8	345	14	B55379 Y177608.r1
34	392	33.7	587	12	BB686319 602638232
35	373	32.1	400	9	A1152313 ud87h02.r
36	373	32.1	441	13	BI967060 1d27h07.y
37	370	31.8	487	10	AM320117 un17b02.x
38	369.5	31.8	698	13	BI906850 603064633
39	354	30.5	260	10	BB596688 BB596688
40	338	29.1	533	10	BB628951 uu30c03.y
41	336	28.9	910	12	BG110063 602279667
42	335	28.8	456	13	BI966255 1e72904.y
43	316	27.2	687	14	B0208433 UI-R-EP0
44	287	24.7	531	13	B1824443 603038693
45	243.5	21.0	639	10	BB642326 BB642326

# ALIGNMENTS

RESULT 1	BF577781	918 bp	mRNA	EST 12-DEC-2000
LOCUS	BF577781			
DEFINITION	602092080F1 NCI_CGAP_C024 Mus musculus cDNA clone IMAGE:4206595 5',			
ACCESSION	BF577781			
VERSION	BF577781.1	GI:11651493		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
	National Institutes of Health, Mammalian Gene Collection (MGC)			





[illegible]

QY	158	LeuValangjiValLeuAlaLeuArgGlySerGluGluGlnPheSerAlaThrAlaIAser	177
Db	374	CTGGTGGATGGGTGCTGCTGGCCCTGGGCTGCTGGAGGAATTTCTACGCCACATCGGGCGAGT	433
QY	178	SerProGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGly	197
Db	434	TTCCTCGGGCCCGACCTCGGCTTCGCCAGGATGTCTGGGCTCTTGCCCTCGGCCACAGG	493
QY	198	SerSerLeuArgGlyLeuArgThrLeuProThrPheAlaHisLeuValAlaAlaProPheLeuThr	217
Db	494	TCTCTCCCTGGCGGATCGGACCCCTCCCTGGGGCCCATCTCAAGGCTGCCCCCTTCTCTCAC	553
QY	218	TyrPheGlyLeuPheGlnValHis	225
Db	554	TACTTCGGACCTCTTCAGATTCTAC	577
RESULT 5			
LOCUS	B1596681	828 bp	linear
DEFINITION	603243254r1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5',		
ACCESSION	B1596681		
VERSION	B1596681.1	GI:15489620	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 828)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@ds-remail.nih.gov		
	Tissue, Procurement: Miklos Palkovits, M.D., Ph.D.		
	cDNA library Preparation: Michael J. Brownstein (NHGRI), Shihai		
	Toshiyuki and Piero Carninci (RIKEN)		
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: L1AM11722 row: k column: 13		
	High quality sequence stop: 776.		
FEATURES	Location/Qualifiers		
source	1..828		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5285892"		
	/clone_lib="NIH_MGC_96"		
	/tissue_type="hypothalamus"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pBluescript (modified		
	pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag		
	); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTNN-3',		
	size-selected for average insert size 2.3 kb and		
	normalized to R0T 5. This is a primary library enriched		
	for full-length clones and constructed using the		
	Cap-trapper method (Carninci, in preparation). Library		
	constructed by M. Brownstein (NIH/NHGRI, National		
	Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	155 a 278 c 223 g 172 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1,10e-80	Length:	828
Score:	861.00	Matches:	165
Percent Similarity:	94.44%	Conservative:	5
Best Local Similarity:	91.67%	Mismatches:	10
Query Match:	74.10%	Indels:	0
DB:	13	Gaps:	0

QY 46 GlutAsnProGlnThrGluSerGlnAspValValProPheLeuGluGlnLeuVal 65  
|||||  
Db 44 GAACGAATCCCGACAGCAAGAAAGCCAGATCCCTGCGCTTCGTAACCGACTACTT 103  
QY 66 ArgProArgSerAlaProLysGlyArgLysAlaAlaProArgArgAlaAlaAla 85  
|||||  
Db 104 CGGCTGCCAAGAGTGCACCTTAAGGCCGGAACACGCGCTCGAAGAGCATCCACCC 163  
QY 86 HisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspLysThr 105  
|||||  
Db 164 CATATGAAATTCATCCACGACCTGACAGACGAGCGACGAGCGAGTGTGACGAGCACA 223  
QY 106 ValSerGlyTyrPglGluThrLysLysLeuSerSerProLeuArgTyrAspArgLys 125  
|||||  
Db 224 TTGAGTGGCTGGAGAGAACCCAGATCAACAGCTCCACCTCTGCTACACCGCCAG 283  
QY 126 IleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPhe 145  
|||||  
Db 284 ATCGGGAGTTTATAGTCAACCGGGCTGGCTCTACTACTGCTACTGCTGACTTT 343  
QY 146 AspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeu 165  
|||||  
Db 344 GATGAGGGAGAGGCTGTCTACCTGAGCTGACTGCTGAGTGTGCTGCTGCTG 403  
QY 166 ArgCysLeuGluGluPheSerAlaThrAlaLysSerProGlyProGlnLeuArgLeu 185  
|||||  
Db 404 CGCTGGCTGGAGATTCACAGCCAGCTGCGACCTTCCTCGGGCCCGACCTCGCCTC 463  
QY 186 CysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeu 205  
|||||  
Db 464 TGCCAGAGTGTCTGGCTGTGGCCCTGCGGCGAGGCTCTCCGTCGATCCGACCTC 523  
QY 206 ProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225  
|||||  
Db 524 CCTGGGCCCATCTCAAGCTGCGCCCTCTCTCCTCCTGAGCTCTCCAGAGTTCAC 583  
RESULT 6  
LOCUS B1766766 834 bp mRNA linear EST 25-SEP-2001  
DEFINITION 60305686F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5206217 5',  
mRNA sequence.  
ACCESSION B1766766  
VERSION B1766766.1 GI:15758344  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 834)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L1AM1517 row: c column: 18  
High quality sequence stop: 772.  
Location/Qualifiers  
1. 834  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5206217"  
/clone\_1db="NIH\_MGC\_122"  
/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;  
Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source

anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 026. Note:  
this is a NIH\_MGC library."   
BASE COUNT 154 a 287 c 223 g 170 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,19e-80 Length: 834  
Score: 861.00 Matches: 172  
Percent Similarity: 92.27% Conservative: 7  
Best Local Similarity: 88.66% Mismatches: 14  
Query Match: 74.10% Indels: 2  
DB: Gaps: 0  
US-09-245-198a-2 (1-225) x B1766766 (1-834)  
QY 33 GlnGluGluLeuThrAlaGluAspArgArgGluProGluLeuAsnProGlnThrGlu 52  
|||||  
Db 3 CAGAGAGAGCTGTGGTGGCAGAGAGACACGAGCCGTGCGAATCCACAGACAGAA 62  
QY 53 GluSerGlnAspValAlaProPheLeuGluGlnLeuValArgProArgArgSerAlaPro 72  
|||||  
Db 63 GAAAGCCAGAGATCTCGCCCTTCTTCAACCGCACTGATGCTGCGCTCGAGAGTGCACCT 122  
QY 73 LysGlyArgLysAlaArgProArgArgAlaIleAlaAlaHisTyrGluValHisProArg 92  
|||||  
Db 123 AAAGCCGGAAGAACACGCGGCTCGAAGAGCATCCAGCCCATTTGAAAGTTCATCCACGA 182  
QY 93 Pro-GlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluThr 112  
|||||  
Db 183 CCGTGGACAGAGCAGGAGCGAGCAGGTGTGACGAGCGGACAGTGTGCTGGAGGAAGC 242  
QY 112 LysLysLeuSerSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrVal 132  
|||||  
Db 243 CAGATATCAACAGCTTCACAGCCCTCTGCGCTACACACCGCCAGATCGGAGAGTTATAGTCAC 302  
QY 132 eArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyr 152  
|||||  
Db 303 CCGGCTGGCTGTCTACTACTGCTGACAGTGTGATGAGGGAAGGCTGTCTA 362  
QY 152 rLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeuGluGluPhe 172  
|||||  
Db 363 CCGTAAGCTGAGCTTGTGTGTGATGTGTGCTGCGCCCTGCGGAGGAATCTTC 422  
QY 172 rAlaThrAlaLysSerSerProGlyProGlnLeuArgLeuGlyGlnValSerGlyLeu 192  
|||||  
Db 423 AGCCACTGCGGCGCACTTCCCTGCGGCGCCAGCTCGCTGTGCCAAGTGTGTGGCTTT 482  
QY 192 uProLeuArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAl 212  
|||||  
Db 483 GGCCCTGGGCGGAGAGTTC-TCCTGTGGGATCCGCAACCTCCCGGGGCCATCTCAAGGC 541  
QY 212 aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225  
|||||  
Db 542 TGCCCTCTTCTCTCACTTCTGAGCTCTTCAGAGTTTCAC 581  
RESULT 7  
LOCUS AM763237 561 bp mRNA linear EST 04-MAY-2000  
DEFINITION ut70009.y1 NCI\_CGAP\_Mem3 Mus musculus cDNA clone IMAGE:315633 5',  
similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS  
; mRNA sequence.  
ACCESSION AM763237  
VERSION AM763237.1 GI:7695174  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 561)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 JOURNAL Tumor Gene Index  
 COMMENT Unpublished (1997)  
 Other\_ESTs: ur70d09.x1  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Lohar Hennighausen Ph.D., Chu Xia Deng Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[image.llnl.gov/image/html/resources.shtml](http://image.llnl.gov/image/html/resources.shtml)

MG1:1058389  
 Seq primer: -40RP from GIBCO  
 High quality sequence stop: 433.  
 Location/Qualifiers  
 1..561  
 /organism="Mus musculus"  
 /strain="129,C57BL/6J,FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3155633"  
 /clone\_lib="NCI-CGAP Mam3"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="10 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: PCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lohar Hennighausen/Chu Xia Deng, NIH  
 Reference for transgenic model: Xu et al., Nature Genetics  
 22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 1 77e-80 Length: 561  
 Score: 857.00 Matches: 167  
 Percent Similarity: 99.40% Conservative: 0  
 Best Local Similarity: 99.40% Mismatches: 1  
 Query Match: 73.75% Indels: 0  
 DB: 10 Gaps: 0

US-09-245-198a-2 (1-225) x AM763237 (1-561)

QY 1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeu 20  
 |||||  
 Db 58 GTGCTGAGCTGGGCGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 117  
 QY 21 GlySerThrPalaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGlu 40  
 |||||  
 Db 118 GGGAGCGGGCAACGCGTGTGCCAGAGACCTTCTCAGAGAGCGTGAACAGAGAGAC 177  
 QY 41 ArgArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPhe 60  
 |||||  
 Db 178 CGCGGGAGAGCCCTGACATGATCCCGACAGAGAGAAAGCCAGATGCGTACCTTTC 237  
 QY 61 LeuGluGlnLeuValAlaArgProArgSerAlaProLysGlyArgLysAlaArgProArg 80  
 |||||  
 Db 238 TTGGACAACACTAGTCCGGCTCGAAGAGTCTCTAAAGCGCGAAGCGCGGCTCGC 297  
 QY 81 ArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAla 100  
 |||||  
 Db 298 CGAGCAATTCACGCCCTTGTGAGTTCATCTCGGCCAGACAGATGAGACACAGCA 357  
 QY 101 GlyValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeu 120  
 |||||  
 Db 358 GGTTGGATGGGACGTGAGTGGCTGGGAAGAGACCAAAATCAACAGCTCCACCCCTCTG 417

QY 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140  
 |||||  
 Db 418 CGCTACGACCGCCAGATGTTGGGATATACATCATCAGGCGGCTCTACTACCTGAC 477  
 QY 141 CysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsn 160  
 |||||  
 Db 478 TGTACAGTCGACCTTGTATGAGGGAAGCGTCTACTCAAGCTGACCTGCTGTAAC 537  
 QY 161 GlyValLeuAlaLeuArgCysLeu 168  
 |||||  
 Db 538 GGTGTGCTGGCCCTGGCTGCTCG 561

RESULT 8  
 LOCUS BQ707185  
 DEFINITION AGENCOURT\_8353983 NIH\_MGC\_113 Homo sapiens CDNA clone IMAGE:6278608  
 5', mRNA sequence.  
 ACCESSION BQ707185  
 VERSION BQ707185.1 GI:21846084  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 948)  
 AUTHORS NIH-MGC <http://imgc.ncbi.nlm.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Mark Watson  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM2466 row: n column: 17  
 High quality sequence start: 24  
 High quality sequence stop: 550.  
 Location/Qualifiers

## FEATURES

source 1..948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6278608"  
 /clone\_lib="NIH-MGC\_113"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH-MGC Library."

BASE COUNT 173 a 334 c 252 g 187 t 2 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 1 9e-77 Length: 948  
 Score: 831.50 Matches: 164  
 Percent Similarity: 92.31% Conservative: 4  
 Best Local Similarity: 90.11% Mismatches: 13  
 Query Match: 71.56% Indels: 2  
 DB: 14 Gaps: 1

US-09-245-198a-2 (1-225) x BQ707185 (1-948)

QY 44 ProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeuGluGln 63  
 |||||  
 Db 3 CCCTCGCAACTGATCCCGACAGACAGAAAGCCAGATCTTCGCTTCGCAACCA 62  
 QY 64 LeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgAlaIle 83



Db	63	CGAGTTGGCCCTCGAAGAAGTCCACCTTAAGCGCGGAATAACAGCGGCTCGAAGCCATC	122
OY	84	AlaAlaHisTyrGluValAlaHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAla	103
Db	123	GCAGCCCATTTATGAAGTTCATCCACAGCTCGACGACGAGCAGCGAGC----	178
OY	104	GlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArgTyrAsp	123
Db	179	GGGACAGTACGTGGCTGGGAGAGACCCAGATCAACAGCTCCAGCCCTCTGGCTACAC	238
OY	124	ArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal	143
Db	239	CGCCAGATCGGGAGATTATAGTCAACCCGGCGTGGCTACTACCTGACTGTCAGG	298
OY	144	HisPheAspGluGlyValAlaValTyrLeuLysIleAsnAspLeuValAsnGlyValLeu	163
Db	239	CACTTTGATGAGGGAGAGGCTCTTACCTCGAAGCTGAGCTGCTGGATGGTGTGCTG	358
OY	164	AlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeu	183
Db	359	GCCCTGGCCCTGGCTGGAGGAATTCACAGCACTGGCGGCAAGTTCCTCGGGCCACGTC	418
OY	184	ArgLeuCysGlnValSerGlyLeuProLeuArgProGlySerSerLeuArgIleArg	203
Db	419	CGCCCTCTCCAGAGTGTCTGGCTGTGTGGCCCTCGCGGACAGGGTCTCTCCCTGGGATCCGC	478
OY	204	ThrLeuProTyrAlaHisIleLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGln	223
Db	479	ACCCCTCCCTGGGCGCATCTCAAGCTCGCCCTTCTCACCCTACTTGGAGACTTCCAG	538
OY	224	ValHis 225	
Db	539	GTTTAC 544	
RESULT 9	BO884231		
LOCUS	BO884231	940 bp	mRNA linear EST 16-AUG-2002
DEFINITION	AGNCOCURT.86622031	Lupsk1_sciatic_nerve	Homo sapiens cDNA clone
ACCESSION	BO884231		
VERSION	BO884231.1	GI:22276239	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 940)		
JOURNAL	NIH-MCIS http://nmc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Dr. James R. Lupski		
	cDNA Library Preparation: Life Technologies, Inc.		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: L1AM13607	row: j	column: 17
	High quality sequence stop: 453.		
FEATURES	Location/Qualifiers		
source	1..940		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_id="IMAGE:6197488"		
	/clone_id="Lupsk1_sciatic_nerve"		
	/sex="male"		
	/tissue_type="sciatic nerve"		
	/dev_stage="adult, 70 yr"		
	/lab_host="DH10B"		
	/note="Vector: pcMW-SPORT6 (Life Technologies); site_1:		

Note: Site-2: SalI: cDNA made by oligo-dT priming. directionally cloned using the following adaptors:  
 5'-TCGACCCACGCGTCG-3' and  
 5'-GACTAGTCTTAGATCGGAGCGGCCGCTT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies.  
 BASE COUNT      170 a      325 c      255 g      190 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.:      5.03e-76      Length:      940  
 Score:      818.00      Matches:      166  
 Percent Similarity:      92.02%      Conservative:      7  
 Best Local Similarity:      88.30%      Mismatches:      11  
                             70.40%      Indels:      4  
 Query Match:      14      Gaps:      0  
 DB:

US-09-245-198A-2	(1-225)	x	BQ884231	(1-940)
OY	42	ArgGIuPProPProGluLeuAsnProGlnInrGluGluSerGlnAspValAlaProPheLeu	61	
Db	1	CAGGACCCGTGGAACTGAATCCCGACAGAAACCAGGATCCGCCCTTCCG	60	
OY	62	GluGlnLeuValArgProArgArGserAlaProLysGlyAlaArgProArgArg	81	
Db	61	AACGCACACTAGTTGGCCTCGCAGAAGTCACCTTAAGGCCGGAACAACGGGCTGAAGA	120	
OY	82	AlaIleAlaIleHisTYrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly	101	
Db	121	GCGATCCGACGCCCATATTGAAGATTTCATCCAGACCTGGACAGAGGAGCGACGACAGT	180	
OY	102	ValAspGlyThrValSerGlyTryPbGlnLunrHyIsIleasnSerSerProLeuArg	121	
Db	181	GTGGACGGGACAGTAGTGCGCTGGAGGAGAACCAATAACAGCTCCAGCCCTTGGCG	240	
OY	122	TyrAspATyGlnIleGlyGluPheThraIleAlaGlyAlaGlyLeuTYrTYrLeuTYrCys	141	
Db	241	TACACGCCCCAGATCGGGGAGATTATTCACACCGCGGCTGGGCTTACTACTGTACTGT	300	
OY	142	GlnValHisPheAspGluGlyLysAlaValTYrLeuLYsLeuAspLeuValAsnGly	161	
Db	301	CAGGTGCACCTTGTGAAGAGGGAGAGCGTGTCACTCAGAGCGTGACTTCTGGTGGATBGT	360	
OY	162	ValIleuAlaLeuArgCysLeuGlnGlnPheSerAlaTrpAlaIleSerSerProGlyPro	181	
Db	361	GTGGCGGCCCTGGCTGCTCGAGGAATTCACGCCATCGCGCCAGTTCCCTCGGGCCC	420	
OY	182	GlnLeuArgLeuCys-GlnValSerGlyLeuLeuProLeuArgProGlySer-SerLeuA	201	
Db	421	CAGCTCGCCCTTCGCCACAGTGTCTGGGGCTGTGGCCCTCGCGCCAGGGGCCCTCCCTGC	480	
OY	201	rGIIeArGYThrLeuProTrpAlaHisLeuLYsAlaAlaI-ProPheLeuThrTYr-PheGl	220	
Db	481	GGATCGCACCCCTCCCTCGGGGCCATTCACAAGGCGCCGCCCTTCTCACTACTTTCGG	540	
OY	220	yLeuPheGlnValHis 225		
Db	541	ACTCTTCAGGTTCC 556		
RESULT 10				
LOCUS	B0671259	963 bp	mRNA	Linear
DEFINITION	AGNCOURT_8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716			
ACCESSION	B0671259			
VERSION	B0671259.1 GI:21782093			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			



PUBMED  
REFERENCE  
AUTHORS

11217851  
5 (bases 1 to 1033)  
Aadachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Butt,C.,  
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,  
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F.,  
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,  
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,  
Numazaki,R., Ohtsu,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,  
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,  
Schirml,L., Shibata,K., Shihata,Y., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,  
Yanaiishi,A., Yoshida,K., Yoshino,M., Muramatsu,M., and  
Hayashizaki,Y.  
Direct Submission  
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suheiho-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGACAGACAGACATCCAGACCTCTTTTTCCTTTTTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through two rounds of normalization to Rot = 20.0 and subtraction  
to Rot = 458.8. Second strand cDNA was prepared with the primer  
adapter of sequence [5'  
GAGACAGACTTTCGAGTAAATTAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved  
with BamHI and XhoI. Vector: a modified plibuescript KS(+) after  
bulk excision from Lambda FLIC I. Cloning sites, 5' end: SalI; 3'  
end: BamHI. Host: DH10B. Retina RNA was provided by Stefano  
Gustcinich (Department of Neurobiology, Harvard Medical School, 220  
Longwood Ave., Boston, MA 02115, USA) whose assistance is  
gratefully acknowledged.

Retina RNA was provided by Stefano Gustcinich (Department of  
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston,  
MA 02115, USA) whose assistance is gratefully acknowledged.

location/Qualifiers

1..1033  
    /organism="Mus musculus"  
    /strain="C57BL/6J"  
    /db\_xref="FANTOM,DB:A930030D13"  
    /db\_xref="MGD:MGI:191867"  
    /db\_xref="taxon:10090"  
    /clone="A930030D13"  
    /tissue\_type="retina"  
    /clone\_id="RIKEN full-length enriched mouse cDNA library"  
    /dev\_stage="adult"  
1..1033  
    /gene="Infsfl2"  
    <1..433  
    /gene="Infsfl2"  
    /note="data source:MGD, source key:MGI:1196259,  
evidence:ISS  
putative  
tumor necrosis factor (ligand) superfamily, member 12"  
    /codon\_start=2  
    /protein\_id="BAB32249.1"  
    /db\_xref="GI:12861641"  
    /db\_xref="MGD:MGI:1196259"  
    /translation="IAAHYEHPRRQDDGAQAGVDGVSGHEEKINSSPRLDRDI  
GEVFVIRAGLTLYCYOVHFBDGRAVYLAKLDLVNGVALPLVEGLFEVSATAASSPEOLR  
LCQVGSLPRPGSSLRIITLPMWHLALAPDLTVFTGLFYOH"

BASE COUNT	228 a	300 c	248 g	256 t	1 others
ORIGIN					
Alignment Scores:					
Pred. NO.:	3.4e-69				1033
Score:	754.00				Matches: 143
Percent Similarity:	100.00%				Conservative: 0
Best Local Similarity:	100.00%				Mismatches: 0
Query Match:	64.89%				Indels: 0
DB:	11				Gaps: 0
US-09-245-198A-2 (1-225) x AK020909 (1-1033)					
OY	83	11ea1aa1ahstyggluvalahisproarprroglyglnaspglyalaglna1aglyal	102		
Db	2	ATTGACGCCCATTAAGAGGTTTCATCTCGGCCAGACAGATGAGCACAAGAGGTGTG	61		
OY	103	ASPGlyThrValSerGlyTfPgluclunhrhrlhys11easnserserserProleuArqTyr	122		
Db	62	GATGGGACAGTGAAGTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCTGGCTTAC	121		
OY	123	ASPARgln11eglygluPheThrVal11ea1ag1aglyleuTyrTyrLeuTyrCysGln	142		
Db	122	GACCCGACGATTGGGGAATTACAGTCAATCAAGGCTGGGCTCTACTACCTGACTGTGAG	181		
OY	143	ValHisPheAspGluGlyLysAlaValTyrLeuTyrLysLeuAspLeuValAsnGlyAl	162		
Db	182	GTGCACCTTGTATGAGGAGAAAGCTCTTCACTCGAAGCTGCACTGCTGTAAGGTTGTG	241		
OY	183	LeuAlaLeuArqCysLeuGluGluPheSerAlaThrAla1aserserProGlyProGln	182		
Db	242	CTGGCCCTGGCGCTGCTGGAAGAATTTCTACGACACAGCAAGCTCTCTGGGCCGACG	301		
OY	183	LeuArqLeuCysGlnValSerGlyLeuLeuProLeuArqProGlySerSerLeuArg11e	202		
Db	302	CTCCGTTTGTGGCAGGTCTGTGGGCTGTGGCGCGTGGGCCAGGGTCTTCCCTTGGATTC	361		
OY	203	ArgThrLeuProTfPAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPhe	222		
Db	362	CGCACCCCTCCCTGGGCGCATCTTAAAGCTGCGCCCTCTTAACCTACTTTGGACTCTTT	421		
OY	223	GlnValHis 225			
Db	422	CAAGTTCAC 430			
RESULT 12					
LOCUS	BM921213	1071 bp	mrna	linear	EST 12-MAR-2002
DEFINITION	AGNCNCOURT_6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752361				
ACCESSION	BM921213				
VERSION	BM921213.1	GI:19371592			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 1071)				
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgaabs-remail.nih.gov				
	Tissue Procurement: Life Technologies, Inc.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	<a href="http://image.lnl.gov">http://image.lnl.gov</a>				
	Plate: LAM12786	row: p	column: 02		
	High quality sequence stop: 656.				
	Location/Qualifiers				



RESULT 14	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
422	AM917574	AM917574	584 bp	EST 25-MAY-2000									
		EST348878 Rat gene index, normalized rat, norvegicus, Bento Soares											
		Rattus norvegicus cDNA clone RGIF49 5' end, mRNA sequence.											
		AM917574											
		AM917574.1 GI:8083328											
		EST.											
		Norway rat.											
		Rattus norvegicus											
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;											
		Rattus.											
		1 (bases 1 to 584)											
		Lee,N.H., Gloder,A., Chandra,I., Mason,T.M., Quackenbush,J.,											
		Kerlanage,A.R. and Adams,M.D.											
		Rat Genome Project: Generation of a Rat EST (REEST) Catalog & Rat											
		Gene Index											
		Unpublished (1998)											
		Contact: Lee, NH											
		The Institute for Genomic Research											
		9712, Medical Center Drive, Rockville, MD 20850, USA											
		Tel: (301)-838-3529											
		Fax: (301)-838-0208											
		Email: nhlee@tigr.org											
		This clone is available through the ATCC, contact the ATCC											
		tel#703-365-2700 for further information											
		Seq primer: M13 Reverse.											
		Location/Qualifiers											
		1..584											
		/organism="Rattus norvegicus"											
		/db_xref="taxon:10116"											
		/clone="RGIF49"											
		/clone_lib="Rat gene index, normalized rat, norvegicus,											
		Bento Soares											
		/tissue_type="mix - brain, ovary, placenta, kidney, lung,											
		liver, embryo, heart, muscle, spleen"											
		/lab_host="SOLR"											
		/note="Vector: pBluescript SK(-); Site.1: EcoRI, site.2:											
		XhoI; Estimated insert size approx.1 kb"											
		XhoI; Estimated insert size approx.1 kb"											
		108 a 188 c 147 g 138 t 3 others											
		BASE COUNT											
		ORIGIN											
		Alignment Scores:											
		Pred. NO.:	9.72e-64	Length:	584								
		Score:	699.										

Db	242	TTCCTAGGCACAGCAGCGAGCTCTCTCCGTGGCCCCGACGCTCCGCTGTGTGCAGTGTCTGG	301
Oy	191	LeuleuPProleuarqPrOgIjSerSeLeuArlglearqThLeuPrioTrPaLaHisLeu	210
Db	302	CTGTGCCTTCGCGGCCAGGGCTCTCCCTTCGGATCGCTACCCTCCCTCGGGCATCTT	361
Oy	211	LysAlaAlaProPhelLeuThrTyRphelLyLeuPheGlnValHis	225
Db	362	AAGCTGCCCCCCTTCCTACTACTTTTGACTTTNTCAAGTTTCAC	406
RESULT 15			
LOCUS	AA870722	445 bp	mRNA linear EST 16-MAR-1998
DEFINITION	vq35g907.r1 Barstead stromal cell line MFLRB8 Mus musculus cDNA		
ACCESSION	AA870722		
VERSION	AA870722.1		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 445)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geiseld,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellensberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HHMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:601556 Seq primer: -28m13 rev2 EF from Amershams High quality sequence svp 420. Location/Qualifiers 1..445 /organism="Mus musculus" /db_xref="taxon:10090" /clone_image="IMAGE:1095324" /clone_lib="Barstead stromal cell line MFLRB8" /cell_line="C2C12 (undifferentiated)" /lab_host="DH10B" /note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker. Site.1: EcoRI; Site.2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACAATCTGAAGTGGAGGCGGCCCTTTTTTTTTTTTTTTTTTTT 3'] ; double-stranded cDNA was ligated to Eco RI adaptors [AATTGCGATCTTGG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV73 vector. Source undifferentiated tissue culture cell line C2C12. Library constructed by Bob Barstead . The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins."		
BASE COUNT	102 a	135 g	89 t
ORIGIN			
Alignment Scores:			
Pred. No.:	7.64e-61	Length:	445
Score:	670.00	Matches:	132
Percent Similarity:	96.38%	Conservative:	1
Best Local Similarity:	95.65%	Mismatches:	5
Query Match:	57.66%	Indels:	1
DB:	9	Gaps:	0



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: March 31, 2003, 07:42:57 ; Search time 11.9352 Seconds

(without alignments)  
781.906 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSTGLALACILGLLVVSL.....PMAHLKAPFLTYGFLFOVH 225

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1162	100.0	225	TN12_MOUSE	O54907 mus musculus
2	1020	87.8	249	TN12_HUMAN	O43508 homo sapien
3	109.5	9.4	272	TNPF5_CHICK	O91868 gallus gall
4	102.5	8.8	316	TN11_MOUSE	O35235 m tumor nec
5	90.5	7.8	260	TNPF5_CANFA	O97626 canis famli
6	90.5	7.8	318	TN11_RAT	O95622 r tumor nec
7	88	7.6	532	PPBN_HUMAN	P10696 homo sapien
8	87	7.5	244	TNPF5_HUMAN	O06643 homo sapien
9	87	7.5	261	TNPF5_AOTTR	O96dm3 aotus trivi
10	87	7.5	261	TNPF5_CALJA	O96dm3 callithrix
11	86.5	7.4	240	TN14_HUMAN	O43557 homo sapien
12	86.5	7.4	1237	B3A2_MOUSE	P13808 mus musculu
13	86	7.4	261	TNPF5_HUMAN	P13965 homo sapien
14	86	7.4	261	TNPF5_MACMU	O93dc7 macaca mula
15	85	7.3	240	TNPF5_MACNE	O96dm7 macaca neme
16	85	7.3	1237	B3A2_RABIT	P48746 oryctolagus
17	84	7.2	278	TNPF6_RAT	P36940 rattus norv
18	83.5	7.2	241	TN13_MOUSE	O94777 mus musculu
19	83.5	7.2	250	TNPF6_MACMU	O94747 macropus eu
20	82	7.1	530	PPB2_HUMAN	P06861 homo sapien
21	82	7.1	535	PPB1_HUMAN	P05187 homo sapien
22	82	7.1	535	PPB3_HUMAN	P05188 homo sapien
23	80.5	6.9	246	C10C_MOUSE	O02105 mus musculu
24	80.5	6.9	1234	B3A2_RAT	P23347 rattus norv
25	79.5	6.8	197	TNPF6_RABIT	P10154 oryctolagus
26	79	6.8	920	PARC_SYNT3	P73077 synechocyst
27	79	6.8	1584	U104_CAEEL	P23678 caenorhabdi
28	79	6.8	2468	MAPB_HUMAN	P46821 homo sapien
29	78.5	6.8	260	TNPF5_PELCA	O97605 felis silve
30	78.5	6.8	1164	PHYD_ARATH	P42497 arabidopsis
31	78	6.7	285	T13B_HUMAN	O92725 homo sapien
32	78	6.7	763	APP2_HUMAN	O06481 homo sapien
33	77.5	6.7	201	TNFB_MACMU	O9xt48 macropus eu

34	77.5	6.7	261	TNPF5_PIG	O95mq5 sus scrofa
35	77.5	6.7	787	NSAP_HUMAN	P49321 homo sapien
36	77	6.6	279	TNPF6_MOUSE	P41047 mus musculu
37	77	6.6	817	NAH1_BOVIN	O28036 bos taurus
38	76.5	6.6	214	SMP_ECOLI	P18838 escherichia
39	76	6.5	695	APP2_MOUSE	O06333 mus musculu
40	75.5	6.5	250	TN13_HUMAN	O75888 homo sapien
41	75.5	6.5	575	MIS_BOVIN	P03972 bos taurus
42	75.5	6.5	999	OXRP_RAT	O63617 rattus norv
43	75.5	6.5	1058	UBA1_HUMAN	P22314 homo sapien
44	75	6.5	1004	SAL2_MOUSE	O9qx96 mus musculu
45	75	6.5	3305	APLP_MANSE	O25490 manduca sex

## ALIGNMENTS

RESULT 1  
TN12\_MOUSE STANDARD: PRT: 225 AA.  
ID TN12\_MOUSE  
AC O54907: O9CTP2: 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (Fragment).  
DE TNFSF12.  
CN Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peritoneal macrophage;  
RC MEDLINE=98070415; PubMed=9405449;  
RA Chicopeurliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H., Hession C., Garcia I., Browning J.L.;  
RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."  
RL J. Biol. Chem. 272:32401-32410(1997).  
RN [2]  
RN SEQUENCE OF 83-225 FROM N.A.  
RP STRAIN=C57BL/6J; TISSUE=Retina;  
RC MEDLINE=21085660; PubMed=11217851;  
RX  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Kono H., Aichi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schirini L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P., Nordone P., Ring B., Ringwald M., Rodriguez N., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Winding L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S., Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
CC -I- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/ADP3. Weak inducer of apoptosis in some cell types. Promotes angiogenesis and the proliferation of endothelial cells. Mediates NF-kappaB activation (By similarity).  
CC -I- SUBUNIT: Homotrimer (potential).  
CC -I- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By similarity).  
CC -I- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
CC -I- PTM: The soluble form is produced from the membrane form by





Query Match 87.8%; Score 1020; DB 1; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 2.6e-84;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 2 LSLGLALACGLLLVSVSGSWATLSAQEPSQELTAEADRREPELNPQTESQDVPFL 61  
 DB 26 LGLGLALACGLLLVSVSGSWATLSAQEPQAEELVAEDQDPSSELNPQTESQDVPFL 85  
 QY 62 EQLVPRRSAPKGRKARRRAIAHYEVHPRGQGAQAGVGYVSGMEETKINSSPLR 121  
 DB 86 NGLVPRRSAPKGRKARRRAIAHYEVHPRGQGAQAGVGYVSGMEETKINSSPLR 145  
 QY 122 YDRQGEFTVIRAGLYLYLCQVHFEDEKAVYKLLDLVNGVALRCLEEFSTAASSPP 181  
 DB 146 YNRQGEFTVIRAGLYLYLCQVHFEDEKAVYKLLDLVNGVALRCLEEFSTAASSPP 205  
 QY 182 QRLCQVSGLLPLRPSSLRIRFLPMALKAAPFLTYGFLQVH 225  
 DB 206 QRLCQVSGLLPLRPSSLRIRFLPMALKAAPFLTYGFLQVH 249

RESULT 3  
 TNF5\_CHICK STANDARD; PRT; 272 AA.

ID TNF5\_CHICK  
 AC Q918D8;  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-l) (CD154 protein).  
 GN TNF5 OR CD40LG OR CD40L.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White leghorn; TISSUE=Spleen;  
 RA Tregaskes C.A., Young J.R., Burnside J.;  
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell proliferation in the absence of co-stimulus as well as IGE production in the presence of IL-4. Involved in immunoglobulin class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an extracellular soluble form (By similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ243435; CAB95748.1; -  
 DR HSSP: P29965; 1ALY.  
 DR InterPro: IPR00363; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KM Cytokine; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 272 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

FT CHAIN 111 272 MEMBER 5, MEMBRANE FORM.  
 FT DOMAIN 1 23 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT TRANSMEM 24 44 MEMBER 5, SOLUBLE FORM (BY SIMILARITY).  
 FT SITE 45 272 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 110 111 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
 FT SITE 190 111 EXTRACELLULAR (POTENTIAL).  
 FT DISULFID 124 124 CLEAVAGE (BY SIMILARITY).  
 FT CARBOHYD 124 124 POTENTIAL.  
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 272 AA; 30862 MW; 5409F24A8E53CCD7 CRC64;

Query Match 9.4%; Score 109.5; DB 1; Length 272;  
 Best Local Similarity 22.5%; Pred. No. 0.012;  
 Matches 58; Conservative 40; Mismatches 107; Indels 53; Gaps 10;

QY 1 VLSGLALACGLLLVSVSGSWATLSAQ-----EPSQELTAEADRREP-----P 45  
 DB 34 VQITGVFLCYLHMKMKMEVLSLNDYITLRVQKQCQTEQDKSTLLDCEKVLKQ 93  
 QY 46 ELNPQTESQDVPFLQVLRPRRSAPKGRK-----APRRRAIAHYEVHPRGQGAQ 100  
 DB 94 DLQCKDRFASBELPKFEMHRSHEHPHLKSRLNETSVAAEERQPIATHLA-----GVKS 145  
 QY 101 GVDGTVSGMEETK-INSSPLRYDRQIGFEFTYIRAGLYLYLCQVHFEDEKA-----V 151  
 DB 146 NTVVRLVLMWTTSYVAPTSLISYHE--GKLWKEKAGLYIYSQVSFCRKAAASAPFLYI 203  
 QY 152 YKLIDLVNGVALRCLEEFSTAASSPPQRLCOV-----SGLLPLRPSSLRIRFLP 206  
 DB 204 YIYLPMEEDRL-MKGLDTHSTSTA-----LCEIQSIRBGVFLRQGDVAVNTD 254  
 QY 207 WAHLKAAPFLTYGFLQV 224  
 DB 255 STAVNVNPGNTFYGMFKL 272

RESULT 4  
 TN11\_MOUSE STANDARD; PRT; 316 AA.

ID TN11\_MOUSE  
 AC Q35235; Q35306; Q9RIY0; Q9JJK8; Q9JJK9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 11 (receptor activator of nuclear factor kappa B ligand) (RANKL) (TNF-related activation-induced cytokine) (TRANCE) (Osteoprotegerin ligand) (OPGL) (Osteoclast differentiation factor) (ODF) (Osteoclastogenesis-inhibitory factor) (OCIF).  
 GN TNFSF11 OR RANKL OR TRANCE OR OPGL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Hydrionia;  
 RX MEDLINE=97460112; PubMed=9312132;  
 RA Wong B.R., Rho J., Arron J., Robinson E., Orlinick J., Chao M., Kalachikov S., Cayani E., Bartlett F.S. III, Frankel W.N., Lee S.Y., Choi Y.;  
 RL "TRANCE is a novel ligand of the tumor necrosis factor receptor family that activates c-Jun N-terminal kinase in T cells.";  
 RL J. Biol. Chem. 272:25190-25194 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Thymic lymphoma;  
 RX MEDLINE=98032977; PubMed=9367155;  
 RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C., Tometsko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D., Galibert L.;

RT "A homologue of the TNF receptor and its ligand enhance T-cell growth  
 RT and dendritic-cell function.";  
 RL Nature 390:175-179(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE-Bone marrow;  
 RX MEDLINE=98227661; PubMed=9568710;  
 RA Lacey D.L., Tims E., Tan H.-L., Kelley M.J., Dunstan C.R.,  
 RA Burgess T., Elliott R., Colombero A., Elliott G., Scully S., Hsu H.,  
 RA Sullivan J., Hawkins N., Davy E., Caparrelli C., Eli A., Qian Y.-X.,  
 RA Kaufman S., Sasse I., Shalhoub V., Senaldi G., Guo J., Delaney J.,  
 RA Boyle W.J.;  
 RT "Osteoprotegerin ligand is a cytokine that regulates osteoclast  
 RT differentiation and activation.";  
 RL Cell 93:165-176(1998).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE-Bone marrow stroma;  
 RX MEDLINE=98188248; PubMed=9520411;  
 RA Yasuda H., Shima N., Nakagawa N., Yamaguchi K., Kinoshita M.,  
 RA Mochizuki S.-I., Tomoyasu A., Yano K., Goto K., Murakami A., Tsuda E.,  
 RA Morinaga T., Higashio K., Udagawa N., Takahashi N., Suda T.;  
 RT "Osteoclast differentiation factor is a ligand for  
 RT osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical  
 RT to TRANCE/RANKL.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3597-3602(1998).  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=129;  
 RX MEDLINE=99214075; PubMed=10196481;  
 RA Kodaira K., Kodaira K., Mizuno A., Yasuda H., Shima N., Murakami A.,  
 RA Ueda M., Higashio K.;  
 RT "Cloning and characterization of the gene encoding mouse osteoclast  
 RT differentiation factor.";  
 RL Gene 230:121-127(1999).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
 RX MEDLINE=21150053; PubMed=11250921;  
 RA Ikeda T., Kasai M., Utsuyama M., Hirokawa K.;  
 RT "Determination of three isoforms of the receptor activator of nuclear  
 RT factor-kappaB ligand and their differential expression in bone and  
 RT thymus.";  
 RL Endocrinology 142:1419-1426(2001).  
 RN [7]  
 RP SEQUENCE OF 139-147, PROCESSING, AND N-GLYCOSYLATION.  
 RX PubMed=10224137;  
 RA Lum L., Wong B.R., Josien R., Becherer J.D., Erdjument-Bromage H.,  
 RA Schindler J., Tempst P., Choi Y., Blobel C.P.;  
 RT "Evidence for a role of a tumor necrosis factor- $\alpha$  (TNF- $\alpha$ )-converting enzyme-like protease in shedding of TRANCE, a  
 RT TNF family member involved in osteoclastogenesis and dendritic cell  
 RT survival.";  
 RL J. Biol. Chem. 274:13613-13618(1999).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 158-316.  
 RX PubMed=11581298;  
 RA Lam J., Nelson C.A., Ross F.P., Teitelbaum S.L., Fremont D.H.;  
 RT "Crystal structure of the TRANCE/RANKL cytokine reveals determinants  
 RT of receptor-ligand specificity.";  
 RL J. Clin. Invest. 108:971-979(2001).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 137-316.  
 RX PubMed=11733492;  
 RA Ito S., Wakabayashi K., Ubukata O., Hayashi S., Okada F., Hata T.;  
 RT "Crystal structure of the extracellular domain of mouse RANK ligand at  
 RT 2.2-Å resolution.";  
 RL J. Biol. Chem. 277:6631-6636(2002).  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF1B/OPG and to  
 CC TNFRSF1A/RANK. Osteoclast differentiation and activation factor.  
 CC Augments the ability of dendritic cells to stimulate naive T-cell  
 CC proliferation. May be an important regulator of interactions  
 CC between T cells and dendritic cells and may play a role in the  
 CC regulation of the T cell-dependent immune response. May also play

CC an important role in enhanced bone-resorption in humoral  
 CC hypercalcemia of malignancy.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted  
 CC (isoforms 1 and 2); Cytoplasmic (isoform 3).  
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THYMUS AND LYMPH NODES BUT  
 CC NOT IN NONLYMPHOID TISSUES AND IS ABUNDANTLY EXPRESSED IN T CELLS  
 CC BUT NOT IN B CELLS. A HIGH LEVEL EXPRESSION IS ALSO SEEN IN THE  
 CC TRABECULAR BONE AND LUNG.  
 CC -1- PTM: N-glycosylated.  
 CC -1- PTM: The soluble form of isoform 1 derives from the membrane form  
 CC by proteolytic processing. The cleavage may be catalyzed by  
 CC ADAM17. A further shorter soluble form was observed.  
 CC -1- DISEASE: DEFICIENCY IN TNFSF1 RESULTS IN FAILURE TO FORM LOBULO-  
 CC ALVEOLAR MAMMARY STRUCTURES DURING PREGNANCY, RESULTING IN DEATH  
 CC OF NEWBORNS. TRANCE-DEFICIENT MICE SHOW SEVERE OSTEOPETROSIS, WITH  
 CC NO OSTEOCLASTS, MARROW SPACES, OR TOOTH ERUPTION, AND EXHIBIT  
 CC PROFOUND GROWTH RETARDATION AT SEVERAL SKELETAL SITES, INCLUDING  
 CC THE LIMBS, SKULL, AND VERTEBRAE AND HAVE MARKED CHONDRODYSPLASIA,  
 CC WITH THICK, IRREGULAR GROWTH PLATES AND A RELATIVE INCREASE IN  
 CC HYPERTROPHIC CHONDROCYTES.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL: AF013170; AAC71061.1; -;  
 DR EMBL: AF019048; AAB66812.1; -;  
 DR EMBL: AF053713; AAC40113.1; -;  
 DR EMBL: AB008426; BAA2425.1; -;  
 DR EMBL: AB022039; BAA36970.1; -;  
 DR EMBL: AB022036; BAA36970.1; JOINED.  
 DR EMBL: AB022037; BAA36970.1; JOINED.  
 DR EMBL: AB022038; BAA36970.1; JOINED.  
 DR EMBL: AB032771; BAA97257.1; -;  
 DR EMBL: AB032772; BAA97258.1; -;  
 DR EMBL: AB036798; BAA97259.1; -;  
 DR PDB: 1J7Z; 12-SEP-01.  
 DR PDB: 1IOA; 13-MAR-02.  
 DR MGI: 1100089; Tnfrsf11.  
 DR InterPro: IPR003636; TNF\_abC.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 DR PROSITE: PS0049; TNF\_2; 1.  
 KW Cytokine; Differentiation; Receptor; Glycoprotein; Transmembrane;  
 KW Signal-anchor; 3D-structure; Alternative splicing.  
 FT CHAIN 1 316  
 FT DOMAIN 139 316  
 FT SITE 1 48  
 FT TRANSMEM 49 69  
 FT DOMAIN 70 316  
 FT SITE 138 139  
 FT CARBOHYD 197 197  
 FT CARBOHYD 262 262  
 FT VARSPPLIC 1 117  
 FT WARSPLIC 14 44  
 FT CONFLICT 99 99 G -> D (IN REF. 2).  
 FT CONFLICT 141 143 MISSING (IN REF. 5).  
 SQ SEQUENCE 316 AA; 34944 MW; 08DF63A2BE00967A CRC64;

Query Match 8.8%; Score 102.5; DB 1; Length 316;  
 Best Local Similarity 24.3%; Pred. No. 0.06; Mismatches 107; Indels 65; Gaps 13;  
 Matches 67; Conservative 37; Mismatches 107; Indels 65; Gaps 13;

QY 3 SIGLALGLLLVVSIGSWATLSAQ-EPSEQELTAEDR-----REPELNPOT 51  
 DB 48 SMFLLALGLGQVVCISIALFLYFRAQMDPNR---ISEDSTHCYFRLRLHFNAGLDOST 104  
 QY 52 EESQDVVF-----FLEQLVRRR--SAPK-----GKARRRRAA 84  
 DB 105 LESEDTLPDSCRMRKMAQOGAVQKELQHVGPFRSGAPAMMEGSLDVAQKKEPAQPF 164  
 QY 85 AHYEHPPPGQDGAQAGVGTGVMEE---TKINSSPLRYDRQIGFETVIRAGLYLY 140  
 DB 165 AHLITNMSIPSGSH---KYTLSSVYHGRGAKISNM-----LSNGKLRVNDQGFYIYL 216  
 QY 141 ---COVHFDEGKAV---YIKLDLV---NGVALRCLSEFSATYASSPQRLI--COV 188  
 DB 217 ANICFRHHETSGVPTDYLQLVVYVYKTSIKIPSSHNLMKGGSTKNMWSGNSSEFHFYSINV 276  
 QY 189 SGLPLRPGSSLRIRTLPMWHLKAPFLTYFGLFOV 224  
 DB 277 GGFELRAGEEISTIOVSNPSLIDPDODATYFGAFKV 312

RESULT 5  
 TNFS\_CANPA  
 ID TNFS\_CANPA STANDARD: PRT: 260 AA.

AC 097626;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand).  
 GN TNFS5 OR CD40LG OR CD40L.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hostle M.H., Willett B.J.;  
 RT "Adjuvant properties of canine CD40L.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell  
 CC proliferation in the absence of co-stimulus as well as IGE  
 CC production in the presence of IL-4. Involved in Immunoglobulin  
 CC class switching (By similarity).  
 CC -1- SUBUNIT: HOMOTRIMER (By similarity).  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN  
 CC EXTRACELLULAR SOLUBLE FORM (By similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF086711; AAD04375.1; -  
 DR HSSP: P29965; IALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF\_1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.

FT CHAIN 1 260 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT FT MEMBER 5, MEMBRANE FORM.  
 FT CHAIN 112 260 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 FT FT MEMBER 5, SOLUBLE FORM (By similarity).  
 FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 23 46 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT FT (POTENTIAL).  
 FT DOMAIN 47 260 EXTRACELLULAR (POTENTIAL).  
 FT SITE 111 112 CLEAVAGE (By similarity).  
 FT DISULFID 177 217 POTENTIAL.  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 260 AA; 28688 MM; 604F69A19E9EB70 CRC64;

Query Match 7.8%; Score 90.5; DB 1; Length 260;  
 Best Local Similarity 28.1%; Pred. No. 0.56; Mismatches 53; Indels 31; Gaps 5;  
 Matches 39; Conservative 16; Mismatches 53; Indels 31; Gaps 5;

QY 111 ETKINSSPLRYDRQIGFET-----VIRAGLYLYLCQVHFDEGKAVYIKL 155  
 DB 128 EASSNPASVLRWAPK-GYTTISSNVLSENGKQLAVKRQGLYVYAQYTFCSNRASSQA 186  
 QY 156 DILVNGVIALRCLSEFSAT-----AASSPQRLCOVS---GLPLRPGSSLRIRTL 205  
 DB 187 PF-----VASLCLHSPGSTERVTLRAASSRGSSKPCGQDSIHLGCVFELHPCGAVFVNVT 241  
 QY 206 PWAHLKAPFLTYFGLFOV 224  
 DB 242 DPSQVSHGTGTFSGFLKL 260

RESULT 6  
 TN11\_RAT  
 ID TN11\_RAT STANDARD: PRT: 318 AA.

AC Q9SE2; Q91Z19;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 11 (receptor activator  
 DE of nuclear factor kappa B ligand) (RANKL) (TNF-related activation-  
 DE induced cytokine) (TRANCE) (osteoprotegerin ligand) (OPGL) (osteoclast  
 DE differentiation factor) (ODF).  
 GN TNFSF11 OR RANKL OR TRANCE OR OPGL.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE= Tibial bone;  
 RX MEDLINE=20540945; PubMed=11092398;  
 RA Xu J.K., Tan J.K., Huang L., Gao X.H., Laird R., Liu D., Wysocki S.,  
 RA Zheng M.H.;  
 RT "Cloning, sequence and functional characterization of the rat  
 RT homologue of receptor activator of NF-kB ligand.";  
 RL J. Bone Miner. Res. 15:2178-2186(2000).  
 RN [2]  
 RP SEQUENCE OF 266-318 FROM N.A.  
 RC STRAIN=Fischer 344;  
 RX PubMed=11804028;  
 RA Odgren P.R., Kim N., van Wesenbeeck L., Mackay C., Mason-Savvas A.,  
 RA Satali F.F., Popoff S.N., Lengner C., van-Hul W., Choi Y.,  
 RA Marks S.C. Jr.;  
 RT "Evidence that the rat osteopetrotic mutation toothless (tl) is not in  
 RT the TNFSF11 (TRANCE, RANKL, ODF, OPGL) gene.";  
 RL Int. J. Dev. Biol. 45:853-859(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFSF11B/OPG and to  
 CC TNFSF11A/RANK. Osteoclast differentiation and activation factor.  
 CC Augments the ability of dendritic cells to stimulate native T-cell  
 CC proliferation. May be an important regulator of interactions  
 CC between T cells and dendritic cells and may play a role in the  
 CC regulation of the T cell-dependent immune response. May also play  
 CC an important role in enhanced bone-resorption in humoral  
 CC hypercalcemia of malignancy.

```

CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in thymus and bone tissues.
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF187319; AAG17031.1; -
DR EMBL: AF425669; AAL23963.1; -
DR HSSP: P50591; 1D0G.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF; 1.
DR ProDom: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF_1; FALSE_NEG.
DR PROSITE: PS00049; TNF_2; 1.
DR Cytokine: Differentiation; Receptor; Glycoprotein; Transmembrane;
KW Signal-anchor.
FT CHAIN 1 318 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT MEMBER 11, MEMBRANE FORM.
FT CHAIN 141 318 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT MEMBER 11, SOLUBLE FORM.
FT DOMAIN 1 47 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 48 68 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 69 318 EXTRACELLULAR (POTENTIAL).
FT SITE 140 141 CLEAVAGE (BY SIMILARITY).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 317 317 I -> M (IN REF. 2).
SQ SEQUENCE 318 AA; 35370 MW; 4B87A4D706AD098F CRC64;

Query Match 7.8%; Score 90.5; DB 1; Length 318;
Best Local Similarity 23.3%; Pred. No. 0.72;
Matches 64; Conservative 36; Mismatches 108; Indels 67; Gaps 12;

QY 6 LALACGLLVVSLGSMATLSAO-EPSOELTAEDR-----REPELNPTEES 54
DB 51 LALLGLGLGVVCSIALFLYFRAMDNR--ISEDSTRFYRLRLRENTGLDSTLES 107
QY 55 QD-----VVPFLDQVLRPRSA-----PKGRKARRRAIAA 85
DB 108 EDTEALPDCSRMKQAFQAGVARELHIVGPRFSGVPAMEGSWLDVARGRPEAQFA 167
QY 86 HVEVHRRPGDGAQAGVDGVSGME---TKNSSPLRLYRQIGETIYRAGLYLYC 141
DB 168 HLTINADIPSGH---KVSLSWYHDRGMAKISNMT---LSNGRLRNQOGFYLYA 219
QY 142 QV-----HFDECK--AVYLKDLLV---NGVLALRCLLEFSATTAASPGQLRL--COVS 189
DB 220 NICEFHHEHSGVPAVDQLQMLVYVVKTSIKIPSSHNLKMGKSTKNMSGNSEPHFYSINVG 279
QY 190 GLLPLRPGSSLRIRTLPMWHLKAPFLTYFGFLQV 224
DB 280 GPFKLKAGEEISVQVSNPSLLDPPDQATYFGARKV 314

RESULT 7
PPRN_HUMAN STANDARD; PRT; 532 AA.
AC P10696; O96CML;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline phosphatase, placental-like precursor (EC 3.1.3.1) (Nagao
DE isozyme) (germ-cell alkaline phosphatase) (PLAP-like) (ALP-1).
GN ALPLP2 OR ALPLP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90283879; PubMed=2162249;
RA Lowe M.E., Straus A.W.;
RT "Expression of a Nagao-type, phosphatidylinositol-glycan anchored
RT alkaline phosphatase in human choriocarcinomas.";
RL Cancer Res. 50:3956-3962(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Choriocarcinoma;
RX MEDLINE=89308696; PubMed=2745460;
RA Watanabe S., Watanabe T., Li W.L., Soong B.-W., Chou J.Y.;
RT "Expression of the germ cell alkaline phosphatase gene in human
RT choriocarcinoma cells.";
RL J. Biol. Chem. 264:12611-12619(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88203632; PubMed=2834730;
RA Millan J.L., Manes T.;
RT "Seminoma-derived Nagao isozyme is encoded by a germ-cell alkaline
RT phosphatase gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:3024-3028(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-157 FROM N.A.
RX MEDLINE=88262578; PubMed=3387245;
RA Shen L.P., Liu H., Kan Y.-W., Kam W.;
RT "5' nucleotide sequence of a putative human placental alkaline
RT phosphatase-like gene.";
RL Nucleic Acids Res. 16:5694-5694(1988).
CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC alcohol + phosphate.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: TRACE AMOUNTS IN THE TESTIS AND THYMUS,
CC AND IN ELEVATED AMOUNTS IN GERM CELL TUMORS.
CC -1- MISCELLANEOUS: IN MOST MAMMALS THERE ARE FOUR DIFFERENT ISOZYMES:
CC PLACENTAL, PLACENTAL-LIKE, INTESTINAL AND TISSUE NON-SPECIFIC
CC (LIVER/BONE/ KIDNEY).
CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X55958; CAA39425.1; -
DR EMBL: J04948; AAA51700.1; -
DR EMBL: J03252; AAA98616.1; -
DR EMBL: BC014139; AAH14139.1; -
DR EMBL: X07247; CAA30232.1; ALU_SEQ.
DR PIR: A31336; A31336.
DR PIR: A34333; A34333.
DR HSSP: P00634; 1ATC.
DR Stena-2DPAGE; P10696; -
DR Genew; HGNC:441; ALPLP2.
DR MIM: 171810; -

```



CC	CARBOND	222	222	N-LINKED (GLCNAC.	) (POTENTIAL).
FT	VASAPLIC	53	77	GLYETADPGQAQOGLEGFLPEE ->	GLGFSRCORRSQ
FT	VARSPLIC	78	244	KIYATPSQSQTDS (IN ISOFORM 2).	
FT	VARIANT	70	70	MISSING (IN ISOFORM 2).	
FT	VARIANT	111	111	/FTID=VAR_013025.	
FT	CONFLICT	60	69	A -> P.	/FTID=VAR_013026.
FT	SEQUENCE	244 AA;	25390 MM;	DPGAQAQGL ->	GLSAPSGSRT (IN REF. 2;
SO				FAL1569459830EDAC CRC64;	
	Query Match	7.5%;	Score 87;	DB 1;	Length 244;
Oy	Best Local Similarity	23.0%;	Pred. No. 1.1;		
Matches	64;	Conservative	26;	Mismatches	80;
				Indels	108;
				Gaps	13;
Oy	3 SIGLALA---	CIGLLLVVSLGSNMTLSAOPSEBELTAEDR-----	-REPEL 47		
Db	18 SLLAVAGATSIVTLT	LAVPIVLAVLVPDGGGLVETFDPAQAQAGLGFKLEE 77			
Oy	48 NPQTESODVVPFLEOLV	PRRSAPKGRKARPAIRAIAAHYEHPPPGDGAQAGVDGVTS 107			
Db	78 EPET----	DLSPGLP-----AAHLTGAPLKGG------	-L 102		
Oy	108 GMEETRINS--SSPLRYDRQIGEFTVIRAGLYELCYOHF-----	-DEGRAVYLKL 155			
Db	103 GWETTKEDAFITSGIQESDAEG-	LALPDGLTYLLCLVGKRAPPGGGDPGGRSVTLRS 161			
Oy	156 DLVNGVALRLCLEBFSATAASSPG-PQLRL-----	-CQVS 189			
Db	162 SL-----	YRAGGAYGCPGTPELLLEGATVTPVLDPARRGYGPLWTVSGRG 208			
Oy	190 GLPLRPSSLRIRILPNANLKAAPFL--	-TYEGLFQV 224			
Db	209 GLVOLRRGERVVYNI--SHPDMDVFARGKTFFGAMV	243			
RESULT 9					
TNF5_AOTTR					
ID TNF5_AOTTR	STANDARD:	PRT:	261 AA.		
AC Q9BDM3:					
DT 15-JUN-2002	(Rel. 41, Created)				
DT 15-JUN-2002	(Rel. 41, Last sequence update)				
DT 15-JUN-2002	(Rel. 41, Last annotation update)				
DE Tumor necrosis factor ligand superfamily member 5	(CD40 ligand) (CD40-				
DE L) (CD154 protein).					
GN TNF5 OR CD40LG OR CD40L.					
OS Acinus trivirgatus (Might monkey) (Douroucouill).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.					
OX NCBI_TaxID=9503;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=Lymphocytes;					
RX MEDLINE=21383618;	PubMed=11491535;				
RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,					
RA Weiss W.R., Ansari A.A.;					
"Cloning, sequencing, and homology analysis of nonhuman primate					
RT Fas/Fas-ligand and co-stimulatory molecules.";					
RT Immunogenetics 53:315-328(2001).					
RL -I- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell					
CC proliferation in the absence of co-stimulus as well as Ige					
CC production in the presence of IL-4. Involved in immunoglobulin					
CC class switching (By similarity).					
CC -I- SUBUNIT: HOMODIMER (By similarity).					
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an					
CC extracellular soluble form (By similarity).					
CC -I- PTM: The soluble form derives from the membrane form by					
CC proteolytic processing (By similarity).					
CC -I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.					
CC -----					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -					

CC	the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	EMBL: AF344860: AAK37542.1: -
DR	HSSP: P29965: 1A1Y.
DR	InterPro: IPR003263: TNF_5.
DR	InterPro: IPR003636: TNF_abc.
DR	InterPro: IPR000478: TNF_family.
DR	Pfam: PF00229: TNF: 1.
DR	ProDom: PD002012: TNF_abc: 1.
DR	ProDom: PD008600: TNF_5: 1.
DR	SMART: SM00207: TNF: 1.
DR	ProSITE: PS00251: TNF_1: 1.
DR	ProSITE: PS50049: TNF_2: 1.
KM	Cycloline; Transmembrane; Glycoprotein; Signal-anchor.
FT	CHAIN 1 261
FT	CHAIN 113 261
FT	DOMAIN 1 22
FT	TRANSMEM 23 43
FT	DOMAIN 44 261
FT	SITE 112 113
FT	DISULFID 178 218
FT	CAROHND 240 240
FT	SEQUENCE 261 AA; 29357 MW; 85E158B507901B5 CRC64;
Query Match	7.5%; Score 87; DB 1; Length 261;
Best Local Similarity	25.98%; Preg. No. 1.2; Mismatches 55; Indels 24; Gaps 6;
Matches 35; Conservative 21;	
Qy	105 TVSGMEE-----TKISSSPLRYDRIGEFYVRAGLYLYLCVHDECKAVYLKLDLVN 160
Db	136 SYLQMAEKGYTMSNNLTLENGKL---TYKROGLYYTNAQVTCNSREASSQAFIAS 192
Qy	161 GVALRCLEEF-----SATAASSPGQLRLC-----QVSGLLPLRPSSSLRITLPMAH 209
Db	193 --LCLKPMPRRFERILLRANHTSSAKP-----CGQGSIHGIGFELQPGASVFVNVTDPQ 246
Qy	210 LKAAPFLYFGLFQV 224
Db	247 VSHGTGFTSPFGLKL 261
RESULT 10	
TNF5_CALJA	STANDARD: PRT; 261 AA.
AC	Q9BDN3:
DR	15-JUN-2002 (Rel. 41, Created)
DR	15-JUN-2002 (Rel. 41, Last sequence update)
DR	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
DE	L) (CD154 protein).
GN	TNFSF5 OR CD40LG OR CD40L.
OS	Callitrix jacchus (Common marmoset).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC	Callitrix.
OX	NCBI_TaxID=9483;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lymphocytes;
RA	MEDLINE=21383618; PubMed=11491535;
RA	Villinger F., Boslik P., Mayne A.E., King C.L., Genain C.P.,
RA	Weiss W.R., Ansari A.A.;
RT	"Cloning, sequencing, and homology analysis of nonhuman primate
RT	Fas/Fas-ligand and co-stimulatory molecules.";
CC	Immunogenetics 53:315-328(2001).
CC	-1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell1







Query Match 7.4%; Score 86.5; DB 1; Length 1237;  
 Best Local Similarity 27.9%; Pred. No. 8.4;  
 Matches 57; Conservative 19; Mismatches 77; Indels 51; Gaps 12;

OY 21 GSWATLSAQEESQELT-AEDRREPELNPQTESQDVVPLEQ-----LVPRRSAPKG 74  
 DB 115 GETPIEGEEDEEERASREGRAPQ-QPSATTPSAVQFLQDECAEKRPETSPSP 173  
 OY 75 RKAPRRRAIAHYEVHPPRGDGAQGV-----DGVSGMEETKI--N 115  
 DB 174 PTQTP-----HQEAAPRASK-GAQTGLVEEMVAVASATAGDGAAGRLTKAOPCH 226  
 OY 116 SSSPLRYROIGEFVVIAGLYLYKCYHFDGKAVYL---KLIDL-----VAGVL-- 163  
 DB 227 RSYNQERRRIGSGMTGVSDA---LIPRVPTDESEAQTATADLDLMSKSHFEDVPGVRRH 283  
 OY 164 ALKCLEFSATMA---SSPGPOLR 184  
 DB 284 LVKRNKAKSTQAAREGREGPTPR 307

RESULT 13  
 ID TNF5\_HUMAN STANDARD; PRT; 261 AA.  
 AC P29965;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-  
 L) (TNF-related activation protein) (TRAP) (T cell antigen Gp39)  
 DE (CD154 antigen).  
 GN TNFSF5 OR CD40LG OR CD40L OR TRAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93076854; PubMed=1280226;  
 RA Graf D., Korthauer U., Mages H.W., Senger G., Kroccek R.A.;  
 RT "Cloning of TRAP, a ligand for CD40 on human T cells.";  
 RL Eur. J. Immunol. 22:3191-3194(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93049181; PubMed=1385114;  
 RA Hollenbaugh D., Grosmaire L.S., Kullas C.D., Chalupny J.N.,  
 RA Briesch-Andersen S., Noelle R.J., Stamenkovic I., Ledbetter J.A.,  
 RA Aruffo A.;  
 RT "The human T cell antigen gp39, a member of the TNF gene family, is a  
 RT ligand for the CD40 receptor: expression of a soluble form of gp39  
 RT with B cell co-stimulatory activity.";  
 RL EMBO J. 11:4313-4321(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A. AND VARIANTS HIGM1 128-ARG-GLY-129 AND PRO-235.  
 RX MEDLINE=93145330; PubMed=7678782;  
 RA Aruffo A., Farrington M., Hollenbaugh D., Li X., Milatovich A.,  
 RA Nonoyama S., Bajorath J., Grosmaire L.S., Stenkamp R., Neubauer M.,  
 RA Roberts R.L., Noelle R.J., Ledbetter J.A., Francke U., Ochs H.D.;  
 RT "The CD40 ligand, gp39, is defective in activated T cells from  
 RT patients with X-linked hyper-IgM syndrome.";  
 RL Cell 72:291-300(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93094757; PubMed=1281209;  
 RA Spriggs M.K., Armitage R.J., Strockbine L., Clifford K.N.,  
 RA Macduff B.M., Sato T.A., Maliszewski C.R., Fanslow W.C.;  
 RT "Recombinant human CD40 ligand stimulates B cell proliferation and  
 RT immunoglobulin E secretion.";  
 RL J. Exp. Med. 176:1543-1550(1992).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93138085; PubMed=7678552;

RA Gauchat J.F.M., Aubry J., Mazzei G.J., Life P., Jomotte T., Elson G.,  
 RA Bonnefoy J.Y.;  
 RT "Human CD40-ligand: molecular cloning, cellular distribution and  
 RT regulation of expression by factors controlling IgE production.";  
 RL FEBS Lett. 315:259-266(1993).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Shimada M., Terasaki H., Ninomiya R., Shimizu S., Nunoi H.,  
 RA Matsuda I.;  
 RT Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 113-117, AND PROCESSING.  
 RX PubMed=8626375;  
 RA Pierrevalle F., Leccanese-Henchoz S., Blasey H., Aubry J.-P., Elson G.,  
 RA Edgerton M.D., Bonnefoy J.-Y., Gauchat J.-F.;  
 RT "Human native soluble CD40L is a biologically active trimer, processed  
 RT inside microsomes.";  
 RL J. Biol. Chem. 271:5965-5967(1996).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 116-261.  
 RX MEDLINE=96131874; PubMed=8589998;  
 RA Karpusas M., Hsu Y.-M., Wang J.-H., Thompson J., Lederman S.,  
 RA Chess L., Thomas D.;  
 RT "2-A crystal structure of an extracellular fragment of human CD40  
 RT ligand.";  
 RL Structure 3:1031-1039(1995).  
 RN [9]  
 RP 3D-STRUCTURE MODELING OF COMPLEX WITH CD40.  
 RX MEDLINE=98266535; PubMed=9605317;  
 RA Singh J., Garber E., van Vlijmen H., Karpusas M., Hsu Y.-M.,  
 RA Zheng Z., Naismith J.H., Thomas D.;  
 RT "The role of polar interactions in the molecular recognition of CD40L  
 RT with its receptor CD40.";  
 RL Protein Sci. 7:1124-1135(1998).  
 RN [10]  
 RP VARIANTS HIGM1 ARG-36 AND GLY-140.  
 RX MEDLINE=93156839; PubMed=7679206;  
 RA Korthauer U., Graf D., Mages H.W., Briere F., Padayachee M.,  
 RA Malcolin S., Ugazio A.G., Notarangelo L.D., Levinsky R.J.,  
 RA Kroccek R.A.;  
 RT "Defective expression of T-cell CD40 ligand causes X-linked  
 RT immunodeficiency with hyper-IgM.";  
 RL Nature 361:539-541(1993).  
 RN [11]  
 RP VARIANT HIGM1 GLU-123.  
 RX MEDLINE=93156840; PubMed=8094231;  
 RA Disanto J.P., Bonnefoy J.Y., Gauchat J.F.M., Fischer A.,  
 RA de Saint Basile G.;  
 RT "CD40 ligand mutations in X-linked immunodeficiency with hyper-IgM.";  
 RL Nature 361:541-543(1993).  
 RN [12]  
 RP VARIANTS HIGM1 PRO-155; ASP-211 AND VAL-227.  
 RX MEDLINE=93174270; PubMed=7679801;  
 RA Allen R.C., Armitage R.J., Conley M.E., Rosenblatt H., Jenkins N.A.,  
 RA Copeland N.G., Bedell M.A., Edelhoff S., Distcheke C.M.,  
 RA Simoneaux D.K., Fanslow W.C., Belmont J.W., Spriggs M.K.;  
 RT "CD40 ligand gene defects responsible for X-linked hyper-IgM  
 RT syndrome.";  
 RL Science 259:990-993(1993).  
 RN [13]  
 RP VARIANTS HIGM1 ALA-126; ARG-140 AND GLU-144.  
 RX MEDLINE=95233438; PubMed=7717401;  
 RA Macchi P., Villa A., Strina D., Sacco M.G., Morali F., Brugnani D.,  
 RA Giliati S., Mantuano E., Fasth A., Andersson B., Zegers B.J.M.,  
 RA Cavaignl G., Reznick I., Levy J., Zan-Bar I., Porat Y., Alto P.,  
 RA Plebani A., Vezroni P., Notarangelo L.D.;  
 RT "Characterization of nine novel mutations in the CD40 ligand gene in  
 RT patients with X-linked hyper IgM syndrome of various ancestry.";  
 RL Am. J. Hum. Genet. 56:898-906(1995).  
 RN [14]  
 RP VARIANTS HIGM1 PRO-155 AND VAL-227, AND VARIANT ARG-219.  
 RX MEDLINE=96133533; PubMed=8550833;  
 RA Lin Q., Rohrer J., Allen R.C., Larche M., Greene J.M., Shigeoka A.O.,

RA Gatti R.A., Derauf D.C., Belmont J.W., Conley M.E.;  
 RT "A single strand conformation polymorphism study of CD40 ligand.  
 RT Efficient mutation analysis and carrier detection for X-linked hyper  
 RT Igm syndrome.";  
 RL J. Clin. Invest. 97:196-201(1996).  
 RN [15]  
 RP VARIANTS HIGM1 ARG-36, CYS-140; SER-231; MET-254 AND GLY-227 DEL.  
 RX MEDLINE=97295077; PubMed=9150729;  
 RA Nonoyama S., Shimadzu M., Toru H., Seyama K., Nunoi H., Neubauer M.,  
 RA Yata J.-I., Ochi H.D.;  
 RT "Mutations of the CD40 ligand gene in 13 Japanese patients with  
 RT X-linked hyper-Igm syndrome.";  
 RL Hum. Genet. 99:624-627(1997).  
 CC -1- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-  
 CC STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.  
 CC INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING.  
 CC -1- SUBUNIT: HOMOTRIMER.  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN  
 CC EXTRACELLULAR SOLUBLE FORM.  
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+  
 CC T-LYMPHOCYTES.  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing.  
 CC -1- DISEASE: DEFECTS IN TNFSF5 ARE THE CAUSE OF AN X-LINKED  
 CC IMMUNODEFICIENCY WITH HYPER-IGM (HIGM1), AN IMMUNOGLOBULIN ISOTYPE  
 CC SWITCH DEFECT CHARACTERIZED BY ELEVATED CONCENTRATIONS OF SERUM  
 CC IGM AND DECREASED AMOUNTS OF ALL OTHER ISOTYPES. AFFECTED MALES  
 CC PRESENT AT AN EARLY AGE (USUALLY WITHIN THE FIRST YEAR OF LIFE)  
 CC RECURRENT BACTERIAL AND OPPORTUNISTIC INFECTIONS, INCLUDING  
 CC PNEUMOCYSTIS CARINI PNEUMONIA AND INTRACTABLE DIARRHEA DUE TO  
 CC CRYPTOSPORIDIUM INFECTION. DESPITE SUBSTITUTION TREATMENT WITH  
 CC INTRAVENOUS IMMUNOGLOBULIN, THE OVERALL PROGNOSIS IS RATHER POOR,  
 CC WITH A DEATH RATE OF ABOUT 10% BEFORE ADOLESCENCE.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- DATABASE: NAME=CD40Lbase;  
 CC NOTE=European CD40L defect database (mutation db);  
 CC WWW=ftp://www.expasy.org/cd40lbase/;  
 CC FTP=ftp://ftp.expasy.org/databases/cd40lbase/;  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD154 entry;  
 CC WWW=ftp://www.ncbi.nlm.nih.gov/prov/cd/cd154.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: X68550; CAA48534.1; -;  
 DR EMBL: Z15017; CAA78737.1; -;  
 DR EMBL: X67878; CAA48077.1; -;  
 DR EMBL: L07414; AAA35662.1; -;  
 DR EMBL: D31797; BAA06599.1; -;  
 DR EMBL: D31793; BAA06598.1; JOINED.  
 DR EMBL: D31794; BAA06599.1; JOINED.  
 DR EMBL: D31795; BAA06599.1; JOINED.  
 DR EMBL: D31796; BAA06599.1; JOINED.  
 DR PIR: S25684; S25684.  
 DR PIR: S26694; S26694.  
 DR PIR: S28017; S28017.  
 DR PIR: S28852; S28852.  
 DR PIR: JH0793; JH0793.  
 DR PDB: 1ALY; 17-SEP-97.  
 DR GeneW: HGNC:11935; TNFSF5.  
 DR MIM: 300386; -;  
 DR MIM: 308230; -;  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.

DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KM Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Antigen;  
 KM Disease mutation; Polymorphism; 3D-structure.  
 FT CHAIN 1 261 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 Query Match 7.4%; Score 86; DB 1; Length 261;  
 Best Local Similarity 25.9%; Pred. No. 1.4;  
 Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;  
 QY 105 TVSGMEP-----TKINSSPLRYDROIGEPYIRAGLYLYCOVHEDGKAVYTKLDLVN 160  
 DB 136 SVLQMAEKGYTWSNNLVTLNGKQL--TVKRGGLYITIAQVFCNSRREASSQAPPIAS 192  
 QY 161 GVIALRCLEEF-----SATAASSPGPOLRLC-----QVSGLLPLRPSSLRITPLPMAN 209  
 DB 193 --LCLKSPGPFERLLRAANTHSSAKP-----CGQSTHLGCVPELQPGASVFNVDPSQ 246  
 QY 210 LKAAPLYTFYGLFQV 224  
 DB 247 VSHGTGFTSFGLLKL 261  
 RESULT 14  
 TNFS\_MACMU  
 ID TNFS\_MACMU STANDARD; PRT; 261 AA.  
 AC Q9BDC7;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-  
 DE L) (CD154 protein).  
 GN TNFSF5 OR CD40LG OR CD40L.  
 OS Macaca mulatta (Rhesus macaque), and  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544; 9531;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND M.MULATTA VARIANTS GLN-60; VAL-204; PRO-206  
 RP AND THR-215.  
 RC SPECIES=M.mulatta, and C.torquatus atys; TISSUE=Lymphocytes;  
 RX MEDLINE=21383618; PubMed=11491535;  
 RA Villingner F., Bostik P., Mayne A.E., King C.L., Genain C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RT "Cloning, sequencing and homology analysis of nonhuman primate  
 RT Fas/FasL ligand and co-stimulatory molecules.";  
 RL Immunogenetics 53:315-328(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell  
 CC proliferation in the absence of co-stimulus as well as Ige  
 CC production in the presence of IL-4. Involved in immunoglobulin  
 CC class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an  
 CC extracellular soluble form (By similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: AF344859; AAK37541.1; -;  
 DR EMBL: AF344841; AAK37600.1; -;  
 DR HSSP: P29965; 1ALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.

```

DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR ProDom: PD002012; TNF_abc; 1.
DR ProDom: PD006600; TNF_5; 1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Polymorphism.
FT CHAIN 1 261
FT CHAIN 113 261
FT CHAIN 113 261
FT DOMAIN 1 22
FT TRANSMEM 23 43
FT TRANSMEM 23 43
FT DOMAIN 44 261
FT SITE 112 113
FT DISULFID 178 218
FT CARBOHYD 240 240
FT VARIANT 60 60
FT VARIANT 204 204
FT VARIANT 206 206
FT VARIANT 215 215
SQ SEQUENCE 261 AA; 29366 MW; AA2E0F1C0F28AB CRC64;

Query Match
Best Local Similarity 7.4%; Score 86; DB 1; Length 261;
Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;

OY 105 TVSGWEE---TKINSSPLRYDRGIEFTVIRAGLYLYCOVHDEGKAVYLIKDLVLN 160
DB 136 SVLQNAEKGYTMSNNLVLENGKQL--TVKROGLYIYAQVTFCSNREASSQAFPIAS 192
OY 161 GVLALRCLEEF-----SATASSPPQRLC-----QVSGLLPLRPGSSIRIRTLPMANH 209
DB 193 --LCUKSPGRFERILLRANHTSSAKP---CGQOSIHLAGVFELOPGASVFVNVTDP SQ 246

OY 210 LKAAPFLTYGFLPOV 224
DB 247 VSHGTGFTSFGILKL 261

RESULT 15
TNF5_MACNE STANDARD: PRT: 240 AA.
AC 09BDM7:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
  l) (CD154 protein).
GN TNFSF5 OR CD40LG OR CD40L.
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RA MEDLINE=21383618; PubMed=11491535;
RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
  Weiss W.R., Ansari A.A.;
RT "Cloning, sequencing, and homology analysis of nonhuman primate
  Fas/Fas ligand and co-stimulatory molecules.";
RL Immunogenetics 53:315-328(2001).
CC -i- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell
  proliferation in the absence of co-stimulus as well as IgE
  production in the presence of IL-4. Involved in immunoglobulin
  class switching (By similarity).
CC -i- SUBUNIT: Homotrimer (By similarity).
CC -i- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an
  extracellular soluble form (By similarity).

```

```

CC -i- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -i- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF344853; AAK37613.1; -.
CC HSSP: P29965; 1ALY.
DR InterPro: IPR003263; TNF_5.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR ProDom: PD002012; TNF_abc; 1.
DR ProDom: PD006600; TNF_5; 1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 240
FT CHAIN 113 240
FT CHAIN 113 240
FT DOMAIN 1 22
FT TRANSMEM 23 43
FT TRANSMEM 23 43
FT DOMAIN 44 240
FT SITE 112 113
FT DISULFID 173 197
FT CARBOHYD 219 219
SQ SEQUENCE 240 AA; 27273 MW; BAE6AA0213CADEC CRC64;

Query Match
Best Local Similarity 7.3%; Score 85; DB 1; Length 240;
Matches 34; Conservative 20; Mismatches 53; Indels 24; Gaps 6;

OY 109 WEE---TKINSSPLRYDRGIEFTVIRAGLYLYCOVHDEGKAVYLIKDLVNGVLA 164
DB 119 WAEKGYTMSNNLVLENGKQL--TVKROGLYIYAQVTFCSNREASSQAFPIAS--LC 173
OY 165 LRCLEEF-----SATASSPPQRLC-----QVSGLLPLRPGSSIRIRTLPMANLRAA 213
DB 174 LKSPGRFERILLRANHTSTKP---CGQOSIHLAGVFELOPGASVFVNVTDP SQVSHG 229
OY 214 PFLTYGFLPOV 224
DB 230 TGFTSFGILKL 240

Search completed: March 31, 2003, 07:58:02
Job time : 18.9352 secs

```

**THIS PAGE BLANK (USPTO)**





RESULT 5  
A:7348  
alpha-glucosidase, probable [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: A87348  
R:Nieman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon  
n, J.; Emlaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete genome sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87348  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-670 <STO>  
A:Cross-references: GB:AE005673; NID:g13422039; PIDN:AMK22781.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC0796

Query Match 7.5%; Score 87.5; DB 2; Length 670;  
Best Local Similarity 24.0%; Pred. No. 10;  
Matches 43; Conservative 24; Mismatches 75; Indels 37; Gaps 6;  
QY 58 VPELEQVLP-----RRSAPKGRKAR-----PRRAIAHYEHVPR----- 93  
DB 498 LPFTRLGPMQYTPGCFRNVA PKDKSFIPTVQTGQALAMVYVDSPLTWSDSP 557  
QY 94 -GQGAQACVD--GTVSGMEETKINSSPLRYDROIGFTVI--RAGLYLYLCOVHDE 147  
DB 558 IYYASPAQLDIFSAVPSWDEIRVLSG-----EIGQYIYARRKGDMMWGAMTND 610  
QY 148 GKAVYLIKDLVNGVIALCLFEFSATASSPGPOLRCQVSGLLPLRGSSLRITRP 206  
DB 611 GTVAVKPSFLGNGAFSAEIRDEGAEPYALKTRTQSVASKDPLTKLAPGGGVIRISP 669

RESULT 6  
A:6066  
lymphotoxin beta - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
C:Accession: A46066  
R:Browning, J.L.; Ngam-ek, A.; Lawton, P.; DeMarinis, J.; Tizard, R.; Chow, E.P.; Hessian  
Cell 72, 847-856, 1993  
A:Title: Lymphotoxin beta, a novel member of the TNF family that forms a heteromeric com  
A:Reference number: A46066; MUID:93208881; PMID:791655  
A:Accession: A46066  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-244 <BRO>  
A:Cross-references: GB:LI1015; NID:g292276; PIDN:AAA36191.1; PID:g292277  
A:Note: sequence extracted from NCBI Backbone (NCBI:128066, NCBI:128067)  
C:Keywords: transmembrane protein

Query Match 7.5%; Score 87; DB 2; Length 244;  
Best Local Similarity 23.0%; Pred. No. 3.7;  
Matches 64; Conservative 26; Mismatches 80; Indels 108; Gaps 13;

QY 3 SIGLALA---CLGLLVVSLGSMATLSAQPESQELTAEDR-----REPEL 47  
DB 18 SLLLVAGATSLVTLIPVILAVLPDQGLVETADPGAQAQGGFGKLPDE 77  
QY 48 NQGTESQDVVPELQVLRPRRSAPKGRKARRRAIAHYEHVPRPGDGAQAGVDGTVS 107  
DB 78 PEP-----DLSPGLP-----AAHLGAPLKGG-----L 102  
QY 108 GMEETKINS--SSPLRYDROIGFTVIRAGLYLYLCOVH-----DEGKAVYIKL 155  
DB 103 GMEETKEDAFILTSQGFSDABE-LALPDQGLYLYLCVLYGKRAPPGGGDPGGRSVTLRS 161  
QY 156 DLLVNGVIALRCLEFSATASSPG-PQLRL-----COVS 189

DB 162 SL-----YRAGAGYCGTPELLELSGAETVTPVLDPARRGCGPLMYTSVGC 208  
QY 190 GLLPLRPGSSLRITLPMHAKAPFL---TYFGLEFQV 224  
DB 209 GLVQLRGGRVYVNI---SHPDVDFARKGTFEGAVMV 243

RESULT 7  
A:83035  
Probable transmembrane sensor PA4895 [imported] - *Pseudomonas aeruginosa* (strain PA01  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 15-Jun-2001  
A:Accession: A83035  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L.  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: A83035  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-340 <STO>  
A:Cross-references: GB:AE004902; GB:AE004091; NID:g9951162; PIDN:AG08280.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4895  
C:Superfamily: *Pseudomonas putida* regulatory protein pupR

Query Match 7.4%; Score 86.5; DB 2; Length 340;  
Best Local Similarity 24.7%; Pred. No. 6;  
Matches 67; Conservative 26; Mismatches 85; Indels 93; Gaps 14;

QY 5 GLALCLGLLV-----VSLGSMATLSAQPESQEL----- 36  
DB 104 GLALLGAGPLVWGMGRQVWRDGCADYLTAVGERRDLYEDGSQVEMNDSALDVRD 163  
QY 37 -----TAEDREPPELNPQTESDVVPFL-----EQLVPRRSAPKGRK 76  
DB 164 AGQRLRLRYGRTYRTADNRREPR-----PFLVTRQGLRALGTFESVR 211  
QY 77 APRRAIAHYE--VHPRPGDGAQAGVDTGVSMEETKINSSPLRYDROIGFTVIRA 134  
DB 212 BEAEVYLAIV-YEGANQVPR--EGASADGRV-----IEGQVRRDRQ-----RI 254  
QY 135 GLYLYLCOVHDEKAVYIKDLVNGVIALCLFEFSATASSPGPOLRCQVSGLLPL 194  
DB 255 G-----PVESASEALAWROGLLVADMDPLR--QWAGELMRYGGESE-CEPS-LDPL 303  
QY 195 RPGSSLRITLPMHAKAPFLTYFGLEFQVH 225  
DB 304 RYSGTFPVDDLP---LALMLAQTHGLRLVH 331

RESULT 8  
A:1789  
band 3-related protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 20-Aug-1999  
C:Accession: A31789  
R:Alper, S.L.; Kopito, R.R.; Libresco, S.M.; Lodish, H.F.  
J. Biol. Chem. 263, 17092-17099, 1988  
A:Title: Cloning and characterization of a murine band 3-related cDNA from kidney and  
A:Reference number: A31789; MUID:89034212; PMID:3182834  
A:Accession: A31789  
A:Molecule type: mRNA  
A:Residues: 1-1237 <ADP>  
A:Cross-references: GB:J04036; NID:g192132; PIDN:AAA65505.1; PID:g309114  
C:Superfamily: band 3 anion transport protein  
C:Keywords: transmembrane protein

Query Match 7.4%; Score 86.5; DB 2; Length 1237;

Best Local Similarity 27.9%; Pred. No. 26; Matches 57; Conservative 19; Mismatches 77; Indels 51; Gaps 12;

QY 21 GSNATLSAQPSOEELT-ADREPELNPQTESODVVFLEQ-----LYRRRSPAPK 74  
 Db 115 GEMPTIEEGEEDSEASGAFRAPO-QSPATPSAVQFLQDEGAEKKEKERTSPSP 173  
 QY 75 RKAPRAIAHYVHPRPDGAQAGV-----DGVSGMEETKI---N 115  
 Db 174 PTPP-----HGEAARPRASR-GAQTGTLYEENVAVASATAGDDGGAAGRPLTKAQPH 226  
 QY 116 SSSPLRYDQIGEFYIRAGLYLYLCOVHDEGKAYL---KLDL-----VNGVL-- 163  
 Db 227 RSYNLQERRRIGSMGTGEQA---LTPRVPTDESAQTATADLDMKSHRFEDVPGVRRH 283  
 QY 164 ALRCLFEFSATA---SSPGPOLR 184  
 Db 284 LVKRNKAGSTQAAREGREGPPRR 307

## RESULT 9

CD40 ligand - human  
 N:Alternate names: glycoprotein 39; hCD40-L protein; T-cell antigen gp39; TRAP protein  
 C:Species: Homo sapiens (man)  
 C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 20-Apr-2001  
 C:Accession: S28017; JH0793; S26694; S28852; S25476; S25684; S30593  
 R:Hollenbaugh, D.; Grosmaire, L.S.; Kullas, C.D.; Chalupny, N.J.; Braesch-Anderesen, S.; EMD 311, 4313-4321, 1992  
 A:Title: The human T cell antigen gp39, a member of the TNF gene family, is a ligand for A:Reference number: S28017; MUID:93049181; PMID:1385114  
 A:Accession: S28017  
 A:Molecule type: mRNA  
 A:Residues: 1-261 <HOL>  
 A:Cross-references: EMBL:Z15017; NID:938483; PIDN:CAA78737.1; PID:938484  
 R:Spr1998, M.K.; Arnltage, R.J.; Strockbine, L.; Clifford, K.N.; Macduff, B.M.; Sato, T. J. Exp. Med. 176, 1543-1550, 1992  
 A:Title: Recombinant human CD40 ligand stimulates B cell proliferation and immunoglobulin A:Reference number: JH0793; MUID:93094757; PMID:1281209  
 A:Accession: JH0793  
 A:Molecule type: mRNA  
 A:Residues: 1-261 <SPR>  
 A:Cross-references: GB:X67878; NID:938411; PIDN:CAA48077.1; PID:938412  
 A:Experimental source: peripheral blood T-cell  
 R:Graf, D.; Korthauer, U.; Mages, H.W.; Senger, G.; Kroccek, R.A. Eur. J. Immunol. 22, 3191-3194, 1992  
 A:Title: Cloning of TRAP, a ligand for CD40 on human T cells.  
 A:Reference number: S26694; MUID:93076854; PMID:1280226  
 A:Accession: S26694  
 A:Molecule type: mRNA  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-261 <GRS>  
 A:Cross-references: EMBL:X68550; NID:937269; PIDN:CAA48554.1; PID:937270  
 R:Gauchat, J.F.; Aubry, J.P.; Mazzel, G.; Life, P.; Jomotte, T.; Elson, G.; Bonnefoy, J. FEBS Lett. 315, 259-266, 1993  
 A:Title: Human CD40-ligand: molecular cloning, cellular distribution and regulation of A:Reference number: S28852; MUID:93138085; PMID:7678552  
 A:Accession: S28852  
 A:Molecule type: mRNA  
 A:Residues: 1-261 <GUA>  
 A:Cross-references: EMBL:L07414; NID:9180123; PIDN:AAA35662.1; PID:9180124  
 A>Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 6-Gln C:Genetics:  
 A:Gene: GDB:CD40LG; HIGM1; IMD3  
 A:Cross-references: GDB:120632; OMIM:308230  
 A:Map position: Xq26-Xq26  
 C:Keywords: glycoprotein; transmembrane protein  
 F:13-44/Domain: transmembrane status predicted <TM>  
 F:45-261/Domain: extracellular status predicted <EXT>  
 F:6,240/Binding site: carbohydrate (Asn) (covalent) status predicted

Query Match 7.4%; Score 86; DB 2; Length 261;  
 Best Local Similarity 25.9%; Pred. No. 4.9;

Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;

QY 105 TVSGWEE-----TKINSSPLRYDROIGEFYIRAGLYLYLCOVHDEGKAYLKLIDLVN 160  
 Db 136 SYLQWAEKGYTMSNNLVLTLENGKOL---TVKRQGLYTYAQTCSNNEASQAFLIS 192  
 QY 161 GVLAIRCLFEF-----SATTAASSPGPOLRLC-----QVSGLLPLRPGSSLRIRLPMWH 209  
 Db 193 --LCLKSPGRFERILLRANTHSSAKP-----CGQGSIHGVEFELQPGASVFNVTDPQ 246  
 QY 210 LKAAPFLTYGFLQV 224  
 Db 247 VSHGTGTSGTLKL 261

## RESULT 10

A56764  
 band 3-related protein, ileum - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 08-Sep-1995 #sequence\_revision 08-Sep-1995 #text\_change 20-Aug-1999  
 C:Accession: A56764  
 R:Chow, A.; Dobbins, J.W.; Aronson, P.S.; Igarashi, P. Am. J. Physiol. 263, G345-G352, 1992  
 A:Title: cDNA cloning and localization of a band 3-related protein from ileum.  
 A:Reference number: A56764; MUID:93035730; PMID:1415547  
 A:Accession: A56764  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1237 <CHO>  
 A:Cross-references: GB:S45791; NID:9256659; PIDN:AAB23488.1; PID:9256660  
 A:Experimental source: New Zealand White rabbit, ileal epithelial cells  
 A>Note: sequence extracted from NCBI backbone (NCBIN:115180, NCBI:115181)  
 C:Superfamily: band 3 anion transport protein

Query Match 7.3%; Score 85; DB 2; Length 1237;

Best Local Similarity 28.1%; Pred. No. 35; Matches 56; Conservative 20; Mismatches 81; Indels 42; Gaps 10;

QY 21 GSNATLSAQPSOEELT-ADREPELNPQTESODVVFLEQLYVPRKSPAPKGRAP 79  
 Db 116 GATPTIEEGEEDSEANENAGARAPTEPSPAPSP-SYQFFLEDGAEKKEKERTSPSP 174  
 QY 80 RRAIAHYVHPRPDGAQAGV-----DGVSGMEETKI---NSSPL 120  
 Db 175 --TLRPHQEAAPV-ATEGAQTGVVEEVAVASGTAGDGASGRPLTKAQPHRSYNL 231  
 QY 121 RYDROIGEFYIRAGLYLYLCOVHDEGKAYL---KLDL-----VNGVLALCLLEE 170  
 Db 232 QERRRIGSMGTGEQA---LTPRVPTDESAQTATADLDMKSHRFEDVPGVRRH 288  
 QY 171 FSATTAASS-----PGPOLR 184  
 Db 289 AKSAGSSRREGREGPPRR 307

## RESULT 11

A49266  
 fas ligand - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
 C:Accession: A49266  
 R:Suda, T.; Takahashi, T.; Golstein, P.; Nagata, S. Cell 75, 1169-1178, 1993  
 A:Title: Molecular cloning and expression of the Fas ligand, a novel member of the A:Reference number: A49266; MUID:94084792; PMID:7505205  
 A:Accession: A49266  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-278 <SDU>  
 A:Cross-references: GB:U03470; NID:9440178; PIDN:AAC52129.1; PID:9440179  
 C:Keywords: glycoprotein; transmembrane protein

Query Match 7.2%; Score 84; DB 2; Length 278;



Best Local Similarity 22.9%; Pred. NO. 7.8;  
Matches 56; Conservative 30; Mismatches 67; Indels 92; Gaps 15;

[illegible]

## RESULT 12

hypothetical protein C35C5.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text-change 18-Feb-2000  
C:Accession: T19756; T21561  
R:White, S.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: Z19173  
A:Accession: T19756  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2962 <W1L>  
A:Cross-references: EMBL:Z78417; PIDN:CAE01693.1; GSPDB:GN00028; CESP:C35C5.1  
A:Experimental source: clone C35C5  
R:Harris, B.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: Z19441  
A:Accession: T21561  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2962 <W1Z>  
A:Cross-references: EMBL:Z78543; PIDN:CAE01756.1; GSPDB:GN00028; CESP:C35C5.1  
A:Experimental source: clone F29G6  
C:Genetics:  
A:Gene: CESP:C35C5.1  
A:Map position: X  
A:Introns: 373/3; 452/2; 525/2; 576/3; 598/3; 782/2; 894/1; 937/3; 1334/3; 1360/3; 1587/3

[illegible]

RESULT 13  
H95947

C:Species: *Sinorhizobium meliloti*  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-sep-2001  
C:Accession: H95947  
R:Ethnan, T.M.; Weldner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Henrich, T.M.; Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing e  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: H95947  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CAC49248.1; PID:915140734; GSPDB:GN00167  
A:Experimental source: strain 1021, megaplasmid pSymB  
R:Gilbert, F.; Ethnan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubb  
L:Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lela  
nebulu, P.; Vandenberg, I.; Vorholter, F.J.; Weldner, S.; Wells, D.H.; Wong, K.; Yeh  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Smb21182  
A:Genome: plasmid  
A:Superfamily: Mycobacterium tuberculosis hypothetical protein RV3272

```

QY      13  LLVAVSLGSM-----ATLSAQEBSOSELTAEDRRE--PEPLNQTEESODVAVPELRQLV  65
Db      11  LELARITAGIPWIGTILADLCADVATKVESPRAGDORTIRGPPVYGEDDERLDAA-YFHACN  69

QY      66  RPRRS-----APYGRKAPRRATAA-----HYE----VHPRP-----  93
      1  : : : : :
Db      70  RGRKRSVLDFTTAAAGCAVRRILAAOSVLLLENFVGLAKYGLDYEESSLKVPNPLIYCSV  129

QY      94  ---GODGA---QAQVDDGVSGWEETKINSSSPLEKYDQIG-EFTVIRAGLY-----YL  139
      1  : : : : :
Db      130  TGFQDQGYAHRAQGYDIYIVQSGMSIMDLTGEPDREPOKIGVAFDITGGLGYAVAAAL  168

QY      140  YCQVHDEGKAVYLUKLDLVNGLVATLRCLEEFSSATAASSPQQLRLCQVSGLLPLRGS  199
      1  : : : : :
Db      190  AQERTGEGQOIIDMALLDCMTGVLANQALNFL-----VSGKAPRRLGN-  233

QY      200  LRITRLPANHAKAPRLIY  218
      1  : : : : :
Db      233  -----AHFNIAVQVE  243

```

RESULT 14

Hypothetical protein SC3A3\_09 - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: T34874  
 R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, August 1999  
 A:Reference number: Z21560  
 A:Accession: T34874  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-467 <SEE>  
 A:Cross-references: EMBL:AL109849; PIDN:CAB52866.1; GSPDB:GN00070; SCOEDB:SC3A3\_09  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC3A3\_09



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 00:47:43 ; Search time 232.957 Seconds  
(without alignments)  
2175.080 Million cell updates/sec

Title: US-09-245-198A-2  
Perfect score: 1162  
Sequence: 1 VLSGLALACILGLLVVSL.....PMAHLKAPFLYFGLFQVH 225

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPO.spool/US09245198/runat\_24032003.163552.25944/app.query.fasta.1.846  
-DB=N\_Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09245198 -QGN\_1.1.338 -tunat\_24032003.163552.25944 -NCPU=6 -ICPU=3  
-NO\_XLPRX -NO\_MMAR -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-MARR.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq.101002.\*  
1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162	100.0	1168	19	AAV18599	Mus musculus tumor
2	1089	93.7	701	20	AAAX23425	Mus TNRL3 DNA.
3	1020	87.8	1236	19	AAV47613	TNF related endoth
4	1020	87.8	1236	22	AAD04350	Human TREPA (TNF r
5	1020	87.8	1353	21	AAA49717	Human PRO207 CDNA
6	1020	87.8	1353	24	ABK40255	CDNA encoding huma
7	1020	87.8	1364	24	ABK34881	Human CDNA encodin
8	1020	87.8	1373	19	AAV18600	Homo sapiens tumor
9	1020	87.8	1421	20	AAV56000	Human tumor necro
10	951	81.8	898	22	AA503964	Expression vector
11	951	81.8	1030	20	AAAX23424	Human TNRL3 DNA.
12	113.5	9.8	2148	24	ABK11680	DNA encoding tumor
13	111	9.6	1221	24	ABK11679	DNA encoding tumor
14	110.5	9.5	978	23	ABL21473	Drosophila melanog
15	106.5	9.2	1630	19	AAV41377	NF-kB receptor act
16	106.5	9.2	1630	19	AAV41377	NF-kB receptor act
17	106.5	9.2	1630	22	AAD15310	Murine receptor ac
18	106.5	9.2	1630	22	AAD08714	Murine receptor ac
19	106.5	9.2	1630	22	AAD05903	Human secreted pro
20	106.5	9.2	3878	22	AAD05230	Human secreted pro
21	106.5	9.2	4034	22	AA564040	Human prosate CDNA
22	106.5	9.2	4034	22	AAH93868	P553S CDNA splice
23	106.5	9.2	4034	24	ABL95411	Human P553S splice
24	106.5	9.2	4894	22	AA564038	Human prosate CDNA
25	106.5	9.2	4894	22	AAH93866	P553S CDNA splice
26	106.5	9.2	4894	24	ABL95409	Human P553S splice
27	106.5	9.2	6976	22	AA564041	Human prosate CDNA
28	106.5	9.2	6976	22	AAH93869	P553S CDNA splice
29	106.5	9.2	6976	24	ABL95412	Human P553S splice
30	105.5	9.1	81800	24	ABK84756	Human CDNA diffe
31	103	8.9	6804	22	AAF82935	HIV-1 gp120 V3 loo
32	102.5	8.8	951	19	AAV69900	Nucleotide sequenc
33	102.5	8.8	951	21	AAA39156	Mouse OSM nucleoti
34	102.5	8.8	951	21	AAZ99965	DNA encoding a mur
35	102.5	8.8	951	21	AAZ49024	Osteoclast formati
36	102.5	8.8	951	21	AAV69886	Nucleic acid encod
37	102.5	8.8	1538	19	ABK11678	DNA encoding tumor
38	102.5	8.8	2191	19	AAV41489	Nucleotide sequenc
39	102.5	8.8	2237	20	AAH80224	Murine TRANCE enco
40	102.5	8.8	2237	24	ABK12880	CDNA encoding mous
41	102.5	8.8	2295	19	AAV70284	Human osteoprotege
42	102.5	8.8	2299	21	AAZ99966	DNA encoding a mur
43	101.5	8.7	7332	24	ABL40768	Constitutively act
44	99.5	8.6	38186	20	AAZ32028	Human MERT1 relate
45	99.5	8.6	38186	22	AAZ90085	AC004449 CDNA clon

## ALIGNMENTS

RESULT 1  
AAV18599  
ID AAV18599 standard; CDNA, 1168 BP.

AAV18599;  
21-JUL-1998 (first entry)

Mus musculus tumor necrosis factor related ligand (TRELL) gene.

TRELL: tumor necrosis factor related ligand; tnfr; treatment;  
cancer; autoimmune disease; immune system; stimulation; suppression;  
graft rejection; ds.

Mus musculus.

OS  
XX  
XX  
FH  
CDS  
Key  
Location/Qualifiers  
2..679

[illegible]

QY	81	ArgAlaIleAlaIleAlaHisIleTyrGluValHisProArgProGlyGlnAspGlyIleAlaIle	100		
Db	242	CGAGCTATTGGACGCCATTATGAGGTTCACTCGGCCGACGAGCATGAGACCAAGCA	301		
QY	101	GlyValAspGlyThrValSerGlyTyrPGLuGluThrLysIleAsnSerSerProLeu	120		
Db	302	GGTGTGATGGGACAGTGAAGTGGCTGGGGAAGACCAACCAATCAACAGCTCCAGCCCTGTG	361		
QY	121	ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr	140		
Db	362	CGCTACGACGCCCGAGATTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACCTGTAC	421		
QY	141	CysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsn	160		
Db	422	TGTCTGGGTCCTTGTATGAGGGAAGGCTGTCTACCTGAAGCTGACTTGGTGTGTAAC	481		
QY	161	GlyValLeuAlaLeuAlaArgCysLeuGluGluPheSerAlaThrAlaAlaSerProGly	180		
Db	482	GGTGTGCTGGCCCTGGCTGCTGGGAAGATTCTCAGCCACAGCAAGCTCTCTCGGG	541		
QY	181	ProGlnLeuAlaGlyLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerLeu	200		
Db	542	CCCCAGCTCCGTTGTGTCAGAGGTGTCGGCTGTGCCCTGGGGCAGGGTTCCTT	601		
QY	201	ArgIleArgThrLeuProTrrPAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGly	220		
Db	602	CGGATCCGACACCTCCCTGGGCTCACTTAAAGCTGCCCTTCTTAACTACTTTGGA	661		
QY	221	LeuPheGlnValHis 225			
Db	662	CTCTTTCAAGTTTCAC 676			
RESULT 2					
ID	AAK23425	AAK23425 standard; DNA: 701 BP.			
AC	AAK23425;				
DT	18-JUN-1999	(first entry)			
XX	Mouse TNRL3 DNA.				
DE					
XX	Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;				
KW	developmental abnormality; gestational abnormality; prostate cancer;				
KW	AP06; AP08; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;				
KW	cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;				
KW	apoptosis; mouse; ss.				
OS	Mus sp.				
XX					
XX					
FH	Key	Location/Qualifiers			
FT	CDS	1..636			
FT		/*tag= a			
FT		/product= "TNRL3"			
XX	W09911791-A2.				
PN	11-MAR-1999.				
XX					
PD	04-SEP-1998; 98WO-US18393.				
PF					
XX	05-SEP-1997; 97US-0924634.				
PR					
XX	(UNIW ) UNIV WASHINGTON.				
PA					
XX					
PI	Chaughary PM;				
XX					
DR	WPI: 1999-205191/17.				
XX					
DR	P-PSDB: AAM93591.				
XX					
PT	New Tumor Necrosis Factor family receptor polypeptides and ligands -				
PT	useful for diagnosis and treatment of prostate cancer and				

PT developmental or gestational abnormalities

XX Example VII; Fig 13B; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
PS receptor polypeptides: ApO4, ApO6, ApO8 and ApO9 or their active  
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
CC their active fragments. ApO4 is useful for diagnosing prostate cancer  
CC by determining levels of ApO4 in an individual. Prostate cancer can also  
CC be treated using ApO4 selective binding agents linked to a therapeutic  
CC moiety. ApO4 polypeptides are also useful for identifying selective  
CC binding agents, useful in diagnosis/treatment of disease by binding of  
CC agents to the polypeptide/active fragment which is extracellular, or  
CC expressed on the cell surface. The binding is preferably performed in  
CC vivo. ApO4 polypeptides/active fragments are also useful for screening  
CC for agonists and antagonists by binding and observing the change in ApO4  
CC activity. Effective pharmacological agents useful in diagnosis or  
CC treatment of disease are also identified using ApO4 polypeptides/active  
CC fragments and ApO4 signal transducer molecules that specifically interact  
CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4  
CC activity. The method is performed in vivo or in vitro. APO polypeptides  
CC are all useful as immunogens for preparing antibodies. ApO4 is also  
CC useful for diagnosis/treatment of developmental or gestational  
CC abnormalities. ApO8 was transfected to human breast carcinoma cell line  
CC MCF-7, and induced apoptosis.

XX Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

#### Alignment Scores:

Pred. No.: 1,94e-90 Length: 701  
Score: 1089.00 Matches: 210  
Percent Similarity: 99.53% Conservative: 0  
Best Local Similarity: 99.53% Mismatches: 1  
Query Match: 93.72% Indels: 0  
DB: Gaps: 0

US-09-245-198A-2 (1-225) x AAX23425 (1-701)

OY 15 LeuValValSerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlu 34  
DB 1 CTGGCTGGTCAAGCTGGAGCTGGGCAACGCTGTGCCACAGACCTTCTAGAG 60  
OY 35 GluLeuThrAlaGluAspArgGluProGluLeuAsnProGlnThrGluGlnSer 54  
DB 61 GAGCTGACAGAGAGAGACGCCCGGAGCCCTGACACTGATCCACACAGAGAAAC 120  
OY 55 GlnAspValProPheLeuGlnGlnLeuValArgProArgGlySerAlaProGlyGly 74  
DB 121 CAGGATGTGTACCTTCTTGGACAACTAGTCCGCGCTCGAAGAGTCTCTAAAGC 180  
OY 75 ArgLysAlaArgProArgAlaAlaAlaHisTyrGlnValHisProArgProGly 94  
DB 181 CGGAAGGGGGCGCTCGCGGAGCTATTCACGCCCAATTAAGAGTTCACTCGGCCAGGA 240  
OY 95 GlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluThrLysIle 114  
DB 241 CAGATGAGACCAACAGAGGTGTGATGACAGTGAAGTGGTGGGAAGACCAAAATC 300  
OY 115 AsnSerSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAla 134  
DB 301 AACAGCTCCAGCCCTCTCGCTACGACCGCCAGATTGGGAATTAAAGTATCAGGGCT 360  
OY 135 GlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLys 154  
DB 361 GGGCTCTACTACCTGTACTGTGAGGTGACATTGTGATGAGGGAAGGCTGTACTCTAGAG 420  
OY 155 LeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThr 174  
DB 421 CTGGACTTGCCTGCTGAAGGCTGTCTGCGCCCTGCGCTGCGGAGAAATCTCACCCACA 480  
OY 175 AlaAlaSerSerProGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeu 194  
DB 481 GCAGCAAGCTCTCTGCGGCCAGCTCCGTTGTGCTCCAGAGTGTCTGGGCTGTGGCGCTG 540

OY 195 ArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaPro 214  
DB 541 CGGCCAGGGCTCTCCCTTCGGATCCGACCCCTCCCTGGCTCATCTTAAGGCTGCCCC 600

OY 215 PheLeuThrTyrPheGlyLeuPheGlnValHis 225  
DB 601 TTCTTAACCTACTTGGACTCTTCAACTGCAC 633

#### RESULT 3

AAV47613  
ID AAV47613 standard; CDNA; 1236 BP.

AC AAV47613;

DT 27-OCT-1998 (first entry)

XX TNF related endothelium proliferative agent gene.

KW ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.

OS Homo sapiens.

FX Key Location/Qualifiers  
FT CDS 1..750  
FT /tag= a  
FT /product= "TREPA"

PN W09835061-A2.

XX 13-AUG-1998.

PD 12-FEB-1998; 98WO-US02859.

XX 10-FEB-1998; 98US-0021706.

PR 12-FEB-1997; 97US-0798692.

PA (ABBO ) ABBOTT LAB.

XX Willey SR;

DR WPI: 1998-447255/38.

XX P-PSDB; AAM29745.

PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
PT treatment of autoimmune disease, tumours and inflammation

PS Claim 11; Page 123-4; 142pp; English.

XX The TNF-related endothelium proliferative agent (TREPA), or its  
CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
CC promote wound healing or tissue grafting, by promoting vascularisation,  
CC also to induce apoptosis for treating cancer and eliminating autoreactive  
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
CC TREPA peptides can also be used to target cytotoxic agents or for  
CC affinity isolation of the corresponding receptor, the nucleic acid for  
CC which can be used to transform tumour cells to render them more  
CC responsive to TREPA and to screen for TREPA mimics.  
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
CC vascularisation), inflammation or a wide range of autoimmune conditions,  
CC conditions involving abnormal stimulation of epithelial cells (e.g.  
CC atherosclerosis), for birth control (inhibiting ovulation and placental  
CC formation) or other angiogenic conditions (e.g. ulcers).

XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

#### Alignment Scores:

Pred. No.: 7.98e-84 Length: 1236  
Score: 1020.00 Matches: 199  
Percent Similarity: 92.86% Conservative: 9  
Best Local Similarity: 88.84% Mismatches: 16

Query Match:	87.78%	Indels:	0
DB:	19	Gaps:	0
US-09-245-198A-2 (1-225) x AAV47613 (1-1236)			
OY	2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly	21	
DB	76 CTGGGCGTGGGCGTGGCGGTGGCTGGCTGGCTGGCGGTGATGTTGGGG	135	
OY	22 SerTPRAlaThrLeuSerAlaInclnProSerGlnGlnLeuThrAlaGlnsparg	41	
DB	136 ACCCGGCAATCTCGCTCCGCCGAGAGGCTGCCGAGAGGACTGGTGCAAGGAGGAC	195	
OY	42 ArgGluProProGluLeuAlaProGlnThrGlnGlnSerGlnAspValValProPheLeu	61	
DB	196 CAGAGACCCGTGGAACTGATCCCAAGACAGAAAGGACAGATCTCGGCTTTCCG	255	
OY	62 GlnGlnLeuValArgProArgArgSerAlaPolyGlyArgLysAlaArgProArgArg	81	
DB	256 AACCACTAGTGTGGCGCTCGAAGAGTGCACTAAAGCCGGGAGAAACAGGGCTCGAAGA	315	
OY	82 AlaIleAlaAlaHisTyrGlnValHisProArgProGlyGlnAspGlyAlaGlnAlaGly	101	
DB	316 GCGATCGCAGCCCATTAATGAAGTTATCATCCAGCACTGGACAGAGAGGACGACAGGT	375	
OY	102 ValAspGlyThrValSerGlyTTPGlnGlnThrLysIleAsnSerSerProLeuArg	121	
DB	376 GTGGAGCGGACAGTAGTGCTGGGAGGAGAACCCAGATCAACAGCTCCAGCCCTGCGCC	435	
OY	122 TyrAspArgGlnIleGlyGlnPheThrValIleArgAlaGlyLeuThrTyrThrTyrCys	141	
DB	436 TACAAACCCGCAATGGGGAGTTTATATAGTCACCCGGGCTGGCTCTACTACTGTACTGT	495	
OY	142 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuLeuValAsnGly	161	
DB	496 CAGGTGCACCTTGATGAGGAGGAAGGCTGTCACTCAAGCTGAGACTGCTGGTGATGT	555	
OY	162 ValLeuAlaLeuAlaArgCysLeuGlnGlnPheSerAlaThrAlaAlaSerSerProGlyPro	181	
DB	556 GTGGTGGCGCTCGCTGGCTGGAGGAATCTCAGCCACGCGCGGCTCCCTGGGCGCC	615	
OY	182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg	201	
DB	616 CAGCTCCGCTGTGCCAGGTGTCTGGGCTGTGGCCCTGGCCAGAGTCTCTCCCTGGCG	675	
OY	202 IleArgThrLeuProTPRAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu	221	
DB	676 ATCCGACACCTCCCTGGGCGCATCTCAAGGCTGCCCTCTCTCAGCACTTCGGACATC	735	
OY	222 PheGlnValHis 225		
DB	736 TTCCAGGTTCAC 747		
RESULT 4			
ID	AAD04350		
XX	AAD04350 standard; cDNA; 1236 BP.		
XX	AAD04350:		
DT	04-JUL-2001 (first entry)		
DE	Human TREPA (TNF related endothelium proliferative agent) cDNA.		
XX			
KW	Human; tumour necrosis factor; TNF; angiogenesis; wound healing;		
KW	TREPA; TNF related endothelium proliferative agent; tumour; metastasis;		
KW	grafting; vulnery; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..750	
FT		/*tag= a	
FT		/product= "Human TREPA (TNF related endothelium	

[illegible]

OY 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
 DB 496 CAGGTGCACTTGTGATGAGGGGAAAGCTGTCTACCTGAAGCTGCACTGTGGATGGT 555  
 OY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181  
 DB 556 GTGCGGGCCCTGCGGTGCTGGAGAAATTCAGCCACTGCGGCGAGTTCCCTCGGGCC 615  
 OY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 616 CAGCTCCCTCTGTCAGAGTGTCTGGCTGGCCCTGCGGCGCAGGGTCTCCCTCGCGG 675  
 OY 202 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221  
 DB 676 ATCCGACCCCTCCCTGCGGCCATCTCAAGGCTGCCCCCTTCCTCACTTCGGACTC 735  
 OY 222 PheGlnValHis 225  
 DB 736 TTCGAGGTTCAC 747

## RESULT 5

AAA49717  
ID AAA49717 standard; cDNA; 1353 BP.

AC AAA49717;

DT 25-SEP-2000 (first entry)

XX Human PRO207 cDNA clone DNA30879-1152.

KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.  
 XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 58..807

FT sig\_peptide 58..177

FT mat\_peptide 178..804

FT /\*tag= c

XX WO200037638-A2.

XX 29-JUN-2000.

XX 02-DEC-1999; 99WO-US28565.

XX 22-DEC-1998; 98US-0113296.

XX 08-MAR-1999; 99WO-US05028.

XX 21-APR-1999; 99US-0130232.

XX 28-APR-1999; 99US-0131445.

XX 14-MAY-1999; 99US-0134287.

XX 20-JUL-1999; 99US-0144758.

XX 26-JUL-1999; 99US-0145698.

XX 15-SEP-1999; 99WO-US21090.

XX 15-SEP-1999; 99WO-US21547.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;

XX Napier MA, Pitti RM, Wood WI;

XX WPI; 2000-442668/38.

XX P-PSDB; AAY95338.

XX Novel composition to inhibit neoplastic cell growth or for treating

XX tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,

XX PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or

PT PRO666 -  
 XX Claim 20; Fig 3; 172pp; English.

CC The present sequence is that of cDNA clone DNA30879-1152  
 CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows  
 CC homology to several members of the tumour necrosis factor family,  
 CC especially human lymphotoxin (23.4%). The cDNA was identified in a  
 CC foetal kidney cDNA library following identification of an expressed  
 CC sequence tag with homology to human Apo-2 ligand. A claimed method  
 CC for inhibiting the growth of a tumour cell comprises exposing the  
 CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,  
 CC PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO666 (see  
 CC AAY95337-49), their agonists or chimeric polypeptides incorporating  
 CC them. The tumour is especially a cancer selected from breast,  
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and  
 CC central nervous system cancer, melanoma and leukaemia. Nucleic  
 CC acids encoding PRO179 etc. are used in the recombinant production  
 CC of the antitumour polypeptides.

XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

## Alignment Scores:

Pred. No.:	8.95e-84	Length:	1353
Score:	1020.00	Matches:	199
Percent Similarity:	92.86%	Conservative:	9
Best Local Similarity:	88.84%	Mismatches:	16
Query Match:	87.78%	Indels:	0
DB:	21	Gaps:	0

US-09-245-198A-2 (1-225) x AAA49717 (1-1353)

OY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21  
 DB 133 CTGGGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 192  
 OY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41  
 DB 193 AGCGGGGATGCGGTGTCCGCCAGAGCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252  
 OY 42 ArgGluProProGluLeuAsnProGlnThrGluSerGlnAspValValProPheLeu 61  
 DB 253 CAGGACCGGTGGAACTGAATCCCGACAGAAAGAAACCGAGATCCCTGCTCTCTG 312  
 OY 62 GlnGlnLeuValArgProArgSerAlaProLysGlyValArgValAlaArgProArgArg 81  
 DB 313 AACGACTAGTTCGGCTCGCAGAGAGTCCACCTTAAGCCCGGAACACGGGCTCGAAGA 372  
 OY 82 AlaIleAlaAlaHisTyrGlnValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101  
 DB 373 GCGATCGCAGCCCATTAATGAAGTTATCCAGCACTGACAGAGAGAGAGAGAGAGAGAGAG 432  
 OY 102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArg 121  
 DB 433 GTGGACGGGACAGTGAAGTGGCTGGGAGAGACCAAGAAACACAGCTCCAGCCCTTGGCG 492  
 OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141  
 DB 493 TACAAACCGCCAGATGGGAGATTATATGACACCGGCTGGGCTCTACTACTGACTACT 552  
 OY 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
 DB 553 CAGGTGCACTTGTGATGAGGGGAAAGCTGTCTACCTGAAGCTGCACTGTGGATGGT 612  
 OY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181  
 DB 613 GTGCGGGCCCTGCGGTGCTGGAGAAATTCAGCCACTGCGGCGAGTTCCCTCGGGCC 672  
 OY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 673 CAGCTCCCTCTGTCAGAGTGTCTGGCTGGCCCTGCGGCGCAGGGTCTCCCTCGCGG 732  
 OY 202 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221

DB 733 ATCCGACCCCTCCCGGGCCATCTCAAGCTGCCCCCTTCTCCTCACTTCTGCGACTC 792  
OY 222 PheGlnValHis 225  
DB 793 TTCACGTTTCAC 804  
RESULT 6  
ABK40255  
ID ABK40255 standard; cDNA; 1353 BP.  
AC ABK40255;  
XX  
XX 15-JUL-2002 (first entry)  
DT  
XX  
XX cDNA encoding human PRO207 polypeptide.  
DE  
XX Human: PRO; benign tumour; malignant tumour; lymphoid malignancy;  
KW Leukemia; neuronal disorder; stromal disorder; blastocoealic disorder;  
KW Inflammatory disorder; immune disorder; angiogenic disorder;  
KW gene therapy; cytostatic; neuroprotective; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200153486-A1.  
PN  
PD 26-JUL-2001.  
XX  
XX 11-FEB-2000; 2000MO-US03565.  
PF  
XX  
XX 08-MAR-1999; 99MO-US05028.  
PR 11-MAR-1999; 99US-123972P.  
PR 11-MAY-1999; 99US-133459P.  
PR 02-JUN-1999; 99MO-US12252.  
PR 22-JUN-1999; 99US-140650P.  
PR 22-JUN-1999; 99US-140653P.  
PR 20-JUL-1999; 99US-144758P.  
PR 26-JUL-1999; 99US-145622P.  
PR 28-JUL-1999; 99US-146222P.  
PR 17-AUG-1999; 99US-149395P.  
PR 31-AUG-1999; 99US-151689P.  
PR 01-SEP-1999; 99MO-US20111.  
PR 15-SEP-1999; 99MO-US21090.  
PR 30-NOV-1999; 99MO-US28313.  
PR 01-DEC-1999; 99MO-US28301.  
PR 01-DEC-1999; 99MO-US28634.  
PR 05-JAN-2000; 2000MO-US00219.  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;  
PI Marsters SA, Pan J, Plitt RM, Roy MA, Smith V, Stone DM;  
PI Matanabe CK, Wood WI.  
XX  
XX WPI; 2002-205567/26.  
DR P-PSDB; AA086129.  
XX  
XX Thirty five nucleic acids encoding PRO polypeptides, useful for  
PT treating benign or malignant tumours, leukemias and lymphoid  
PT malignancies, inflammatory, angiogenic and immunologic disorders -  
XX  
XX Claim 50; Fig 3; 302pp; English.  
XX  
XX The present invention relates to the isolation of novel human PRO  
CC polypeptides and the polynucleotide sequences encoding them. The  
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are  
CC useful for treating benign or malignant tumours (e.g. renal, kidney,  
CC bladder, breast, etc), leukemias and lymphoid malignancies, other  
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,  
CC macrophagal, stromal and blastocoealic disorders, inflammatory, immune  
CC and angiogenic disorders. The polynucleotide sequences are also  
CC useful in gene therapy. ABK40254-ABK40288 encode for the human PRO  
CC polypeptides of the invention.

XX  
SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;  
Alignment Scores:  
Pred. No.: 8 95e-84 Length: 1353  
Score: 1020.00 Matches: 199  
Percent Similarity: 92.86% Conservative: 9  
Best Local Similarity: 88.84% Mismatches: 16  
Query Match: 87.78% Indels: 0  
DB: 24 Gaps: 0  
US-09-245-198A-2 (1-225) x ABK40255 (1-1353)  
OY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21  
DB 133 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 192  
OY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41  
DB 193 AGCGGCGCATCGCTGTCGCCCGAGAGCTGCGCCAGAGAGCTGCTGGCAGAGAGAGC 252  
OY 42 ArgGluProProGluLeuAsnProGlnThrGluSerGlnAspValValProPheLeu 61  
DB 253 CAGGACCCGTCGGAACCTGAATCCCGACAGAGAAAGCCAGGATCTCGGCTTCTCTG 312  
OY 62 GlnGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArg 81  
DB 313 AACCGACTAGTTGGGCTGCGCCAGACAGTGCACCTAAAGGCGGAAACAGCGGCTGAGA 372  
OY 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspLysAlaGlnAlaGly 101  
DB 373 GCGATCGCAGCGCCATATGATGATTCACCGACCTGCGACGACGACGACGACGACGAG 432  
OY 102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArg 121  
DB 433 GTGACGGGACAGTGAATGCTGGAGAGAACAGACATCACAGCTTCAGCCCTTGCGG 492  
OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141  
DB 493 TACACCGCCGACGATCGGGAGGTTATGATCACCGCGGGGCGGCTGACTACTGCTG 552  
OY 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
DB 553 CAGGTCCACTTGTATGATGAGGAGGCTGCTTACCTAAGCTGACCTTGCTGGTGTGTGT 612  
OY 162 ValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaIleSerSerProGlyPro 181  
DB 613 GTGCTGGCCCTGCGCTGCTGCGAGGAAATTCACGCCACTGCGGCGAGTTCCCTCGG 672  
OY 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
DB 673 CAGCTCGGCTTGCAGGAGTGTGCGGCTGCTGCGGCTGCGGCGGAGGCTCTCCGCG 732  
OY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221  
DB 733 ATCCGACCCCTCCCGGGCCATCTCAAGCTGCCCCCTTCTCCTCACTTCTGCGACTC 792  
OY 222 PheGlnValHis 225  
DB 793 TTCACGTTTCAC 804  
RESULT 7  
ABK34881  
ID ABK34881 standard; cDNA; 1364 BP.  
XX  
XX  
AC ABK34881;  
XX  
XX 08-MAY-2002 (first entry)  
DT  
XX  
XX Human cDNA encoding secreted protein #19.  
DE  
XX Human: secreted protein; gene; ss; nutritional supplement; haemophilia;  
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;



KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;  
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
 KW tissue regeneration; wound healing; burn; haematopoiesis;  
 KW myeloid cell deficiency; lymphoid cell deficiency.  
 OS  
 XX Homo sapiens.  
 XX MO200177288-A2.  
 XX PD 18-OCT-2001.  
 XX PF 29-MAR-2001; 2001WO-US10224.  
 XX PR 06-APR-2000; 2000US-195582P.  
 XX PA (GEMV ) GENETICS INST INC.  
 XX PI Mong GG, Clark HF, Fachtel K, Agostino MJ, Howes SH, Resnick RJ;  
 PI Guluketa K, Graham JR;  
 DR WPI; 2002-179321/23.  
 XX PF Five hundred and ninety two polynucleotides derived from a variety of  
 PT human tissue sources which encode secreted proteins, useful for  
 PT treating immune deficiencies and disorders such as autoimmune disorders  
 PS  
 XX Claim 1; Page 82; 372pp; English.  
 XX The invention relates to 592 polynucleotides which have been derived from  
 CC a variety of human tissue sources and which encode novel secreted  
 CC proteins. The polynucleotides can be used as probes for the  
 CC identification and isolation of full length cDNA and genomic DNA. The  
 CC polynucleotides and proteins can also be used as nutritional supplements.  
 CC The proteins are useful in the treatment of various immune deficiencies  
 CC and disorders such as viral infections, bacterial infections, fungal  
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
 CC and conditions (e.g. asthma). They are also useful for treating  
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
 CC useful for tissue regeneration, for wound healing and in the treatment of  
 CC burns, incisions and ulcers. The proteins are also useful for regulating  
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
 CC Sequences ABR34863-ABR35454 represent polynucleotides of the invention.  
 XX  
 SQ Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 9.04e-84 Length: 1364  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 87.78% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-245-198a-2 (1-225) x ABR34881 (1-1364)  
 QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21  
 DB 161 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 220  
 QY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGlnGluLeuThrAlaGluAspArg 41  
 DB 221 AGCGGCGCATGCTGCTGCCGAGAGCGCTGCCGAGAGGAGGAGGAGGAGAC 280  
 QY 42 ArgGluProPArgGluLeuAsnProGlnThrGlnGluSerGlnAspValValProPheLeu 61  
 DB 281 CAGGACCGCTGGAGTGAATGCCAGACAGAAAGCAGATCTGCGCTTCTCTG 340

QY 62 GluGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArgArg 81  
 DB 341 AACGCACTAGTTCGGCTTCGCGAGAGTCCACCTAAAGCCGGAAACACGGGCTCGAAGA 400  
 QY 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101  
 DB 401 GCGATCGCAGCCCATTTATGAAGTTCATCCACGACTGACAGAGGAGGAGGAGGAGGAGT 460  
 QY 102 ValAspGlyThrValSerGlyTyrGlnGluThrLysIleAsnSerSerSerProLeuArg 121  
 DB 461 GTGGACGGGACAGTGAAGTGGCTGGAGAGACCAAGAAACACAGCTCCAGCCCTCTGGCC 520  
 QY 122 TyrAspArgGlnIleGlyLuphethrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141  
 DB 521 TACAAACGCCCATGATCGGGAGATTATAGTACACCGGGCTGGGCTTACTACTGATGCT 580  
 QY 142 GlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuLeuValAsnGly 161  
 DB 581 CAGTGCACCTTTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 640  
 QY 162 ValLeuAlaLeuAlaArgCysLeuGlnGluLuphethrValIleAlaIleAspSerProGlyPro 181  
 DB 641 GTGCTGGCCCTGCGCTGCTGAGAAATCTCAGCCACTGCGGAGTTCCTCGGCGCC 700  
 QY 182 GlnLeuAlaGlyLeuValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 701 CAGCTCCGCTCTGCTGCGAGGAGTGTGGCTGTGGGCTTGGCGGCTCTCTCTGCGG 760  
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221  
 DB 761 ATCCGACCCCTCCCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 820  
 QY 222 PheGlnValHis 225  
 DB 821 TTCAGGTTTCAC 832  
 RESULT 8  
 ID AAV18600 standard; cDNA; 1373 BP.  
 AC AAV18600;  
 XX 21-JUL-1998 (first entry)  
 XX DE Homo sapiens tumour necrosis factor related ligand (TRELL) gene.  
 XX KW TRELL; tumour necrosis factor related ligand; tnfr; treatment;  
 KW cancer; autoimmune disease; immune system; stimulation; suppression;  
 KW graft rejection; ds.  
 XX OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH 1..852 /\*tag= a  
 FT /\*note= "tumour necrosis factor related ligand"  
 FT CDS  
 PN MO9805783-A1.  
 XX 12-FEB-1998.  
 XX PD 07-AUG-1997; 97WO-US13945.  
 XX PF 18-MAR-1997; 97US-0040820.  
 XX PR 07-AUG-1996; 96US-0023541.  
 XX PR 18-OCT-1996; 96US-0028515.  
 XX PA (BIOJ ) BIOGEN INC.  
 XX PA (UTGE-) UNIV GENEVA FACULTY MEDICINE.  
 XX PI Browning JL, Chicheportliche Y;  
 XX WPI; 1998-145619/13.

DR P-PSDB; AAW47525.

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
PS Claim 2; Pages 48-50; 69pp; English.

XX The sequence is that encoding human tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon- gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. The DNA sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of prepare probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.  
XX

SQ Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;

Alignment Scores:

Pred. No.:	9,12e-84	Length:	1373
Score:	1020.00	Matches:	199
Percent Similarity:	92.86%	Conservative:	9
Best Local Similarity:	88.84%	Mismatches:	16
Query Match:	87.78%	Indels:	0
DB:	19	Gaps:	0

US-09-245-198A-2 (1-225) x AAV18600 (1-1373)

QY 2 GlnSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValValSerLeuGly 21  
DB 181 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 240  
QY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluGluLeuThrAlaGluAspArg 41  
DB 241 AGCGGGCGATCGCTGCCGCCAGAGCCCTGCCAGAGCAGCTGGCGGAGAGAGAG 300  
QY 42 ArgGluProGluLeuAspProGlnThrGluGluSerGlnAspValValProPheLeu 61  
DB 301 CAGGACCCGTCGAGACTGAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
QY 62 GlnGluLeuValArgProArgSerAlaProGlyValArgGlyValArgProArgArg 81  
DB 361 AACCGACTACTTCGGGCTGCAGAGTGCACCTAAAGGCGGAAACACGCGGCTCGAGA 420  
QY 82 AlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 101  
DB 421 GCATGCGACCCCATTAATGAATTCACGACCTGGAGAGAGAGAGAGAGAGAGAGAG 480  
QY 102 ValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeuArg 121  
DB 481 GTGGAGCGGACACTGACTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
QY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuValTyrTyrCys 141  
DB 541 TACACCGCCGAGATCGGGAGAGTTATAGTACACCGGCTGGGCTGACTACTACTACT 600  
QY 142 GluValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
DB 601 CAGGTGACTTTGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
QY 162 ValGluAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerSerProGlyPro 181

DB 661 GTGCTGGCCCTGGCTGGCTGGAGGAATTCAGCCACTGCGGCGCAGTCCCTGGGCCC 720  
QY 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
DB 721 CAGCTCCGCTCTGCGCAGGTGTGTGGGCTGTGGGCTGGGCTGGGCTGGGCTGGG 780  
QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221  
DB 781 ATCGGACCCCTCCCTGGGCGCATTCAGAGCTGCCCTCTCTCAGCTACTGAGACTC 840  
QY 222 PheGlnValHis 225  
DB 841 TTCAGAGTTCAC 852

RESULT 9

AA556000 standard; DNA; 1421 BP.

XX AAX56000;  
AC AAX56000;  
XX  
DT 15-JUL-1999 (first entry)  
XX  
DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.  
XX  
KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;  
KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
KW cancer; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 92..841  
FT /tag= a  
FT /product= "Apo-3 ligand"  
XX  
PN WO919490-A1.  
XX  
PD 22-APR-1999.  
XX  
PF 09-OCT-1998; 98WO-US21407.  
XX  
PR 17-DEC-1997; 97US-0069862.  
PR 10-OCT-1997; 97US-0062037.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Marsters SA, Pittl R;  
XX  
DR WPI; 1999-287982/24.  
DR P-PSDB; AAT09369.  
XX  
PT New human Apo-3 ligand (a tumour necrosis factor) homologue  
XX  
PS Claim 18; Fig 1; 74pp; English.  
XX

CC The present sequence encodes a human tumour necrosis factor (TNF) and  
CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
CC cytosolic activity. Apo-3 ligand can be used to induce apoptosis in  
CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and  
CC to induce JNK/SAPK-dependent responses in mammalian cells.

SQ Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;

Alignment Scores:			
Pred. No.:	9.53e-84	Length:	1421
Score:	1020.00	Matches:	199
Percent Similarity:	92.86%	Conservative:	9
Best Local Similarity:	88.84%	Mismatches:	16
Query Match:	87.78%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198A-2 (1-225) x AAX56000 (1-1421)

OY	2	LeSerTlaugLYleuAlaleuAlaCysIeuGluLYleuIleuIleuValValValSerIeuGly	21
Dd	167	CTGGCCCTGGCGCTGGCGCTGGCGCTGGCGCTTCCTGGCCGTGCAGTTTGGGG	226
OY	22	SerTrPalaThrLheuSerAlaGlnIupProSerGIngluGluLheuThrAlaIupSParg	41
Dd	227	AGCGGGCATGGCTGTCCGCCAGGAAGCTGCCAGAGAGACTGTGTGCACAGAGAGAC	286
OY	42	ArgGIuProPGroLYleuAsnProGlnThrGluGluSerGlnASPValAlPropheIeu	61
Dd	287	CAGGACCCTGGCAAGTCAATCCCCAGACAGAAGAAAGCACAGATCCCTGCTTCCTG	346
OY	62	GluGlnIeuValArgProIargrGrserAlaProlysGIArgLYsAlaArgProIargr	81
Dd	347	AACCCAGTAGTTCGGCCCTCCGCAAGATGTCACTAAAGCCGGAAAAACAGCGCTCGA	406
OY	82	AlaIleAlaAlaHisTYrGluValAlaHSProlArgProGlyGlnASPglVALagIlnaIagly	101
Dd	407	GGGATCGGACGCCATTATGTAAGTTTATCCACAGACTGGACAGAGGAGCGCAGCGAGT	466
OY	102	ValASPglTYrThValSerGIYrPGLuGluthrLYsIleAsnSerSerProIeuArg	121
Dd	467	GTGGACGGGACAGTAGTGCTGGGAGGAGAACCAAAATCAACAGCTCCAGCCCTCGCC	526
OY	122	TyrASpArgGlnIleGlyGluPhetherValIleArgAlaGlyLeuTYrTYrIeuTYrCys	141
Dd	527	TACAACCCCGCAATGGGGAGTTTATAGTACACCGGGCTGGCTCTACTACCTGTACTGT	586
OY	142	GlnValHisPheASPglUGLYsAlaValTYrIleuLYsIleASPleuIeuValasngly	161
Dd	587	CAGGTGCACATTGTGATGAGGGAGAAAGCTGTCTACTGAACTGTGACTTGTGTGATGT	646
OY	162	ValIeuAlaIeuarGryCysIeuGluIupheseralarThraAlaIaSerSerProGIPro	181
Dd	647	GTGCTGGCCCTGGCTGCCTGGAGGAATCTCAGACACAGCGGCGATTCCCTCGGGCCC	706
OY	182	GlnIeuArGLeuCysGlnValSerGIYleuIeuProIeuArgProGlySerSerIeuArg	201
Dd	707	CAGCTCCGCTCTGCAGGTGTCTGGGCTGTGGCCCTCGGCCCAAGGGTCTCCCTCGCGG	766
OY	202	IleArGYrIleProTrPalaHisIleuLYsAlaIaProphelueThrTYrphegIyeu	221
Dd	767	ATCCGACACCTCTCCCTGGGCCCATCTCAAGGCTGCCCTTCCTCACTTCCGACTC	826
OY	222	pheGlnValHis 225	
Dd	827	TTCCAGGTTTCAC 838	
<b>RESULT 10</b>			
ID	AAS03964	standard; DNA; 898 BP.	
XX	AAS03964;		
AC			
DT	26-SEP-2001	(first entry)	
XX			
DE	Expression vector pDC409-IL2-TWEAK fusion protein-encoding DNA.		
XX			
KW	TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;		
KM	ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;		
KW	retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;		
KV	rubeosis; uveitis; macular degeneration; arthritis; rheumatism; ds;		
KW	corneal graft neovascularisation; psoriasis; metastatic condition;		
KV	malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;		
KW	preneoplastic condition; myocardial angioneutis; wound granulation;		
KM	scleroderma; vascular adhesion; telangiectasia; ischemia; human;		
KV	atherosclerotic plaque neovascularisation; coronary atherosclerosis;		
KW	peripheral atherosclerosis; pDC409-IL2-TWEAK; TWEAK receptor; TWEAKR;		
XX	fusion protein.		
OS	Homo sapiens.		
CS	Synthetic.		
XX			

```

FH   Key      Location/Qualifiers
FT   CDS      52..873
FT           /*tag= a
FT           /product= "Fusion protein comprising a growth hormone
FT           leader, a leucine zipper multimerisation
FT           domain, and human TWEAK extracellular
FT           domain"
XX
XX      WO200145730-A2.
XX
XX      28-JUN-2001.
XX
XX      19-DEC-2000; 2000WO-US34755.
XX
XX      20-DEC-1999; 9905-0172878.
XX      10-MAY-2000; 2000US-0203347.
XX
XX      (IMMV ) IMMUNEX CORP.
XX
XX      Wiley SR;
XX
XX      WPI; 2001-417975/44.
XX      P-PSDB; AAU03499.
XX
XX      Modulating angiogenesis in a mammal for treating diseases mediated by
XX      angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX      peripheral tissue, by administering antagonist or agonist of TWEAK
XX      receptor
XX
XX      Example 1; Page 39-40; 46pp; English.
XX
XX      The sequence represents a DNA from the expression vector
XX      pDC409-L2-TWEAK, which encodes a fusion protein comprising a growth
XX      hormone leader, a leucine zipper multimerisation domain, and the
XX      extracellular domain of human TWEAK. The fusion protein was used in
XX      the isolation of human TWEAK receptor (TWEAKR)-expressing clones
XX      from a COS cell human cDNA library. The TWEAK protein is a
XX      member of the tumour necrosis factor (TNF) family and induces
XX      angiogenesis. TWEAKR may therefore be used to screen for and develop
XX      TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
XX      used in the treatment and diagnosis of human disease. The disorders
XX      mediated by angiogenesis include ocular disorders characterised by ocular
XX      neovascularisation such as diabetic retinopathy, neovascular glaucoma,
XX      retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
XX      rubeosis, uveitis, macular degeneration and corneal graft
XX      neovascularisation, and inflammatory diseases such as arthritis,
XX      rheumatism and psoriasis. Other treatable diseases include malignant and
XX      CC metastatic conditions such as sarcomas and carcinomas, benign tumours and
XX      CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
XX      CC scleroderma, vascular adhesions, atherosclerotic plaque
XX      CC neovascularisation, telangiectasia, wound granulation, coronary
XX      CC atherosclerosis, peripheral atherosclerosis and ischaemia.
XX
XX      SQ      Sequence 898 Bp; 187 A; 266 C; 267 G; 178 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.:      1,06e-77      Length:      898
XX      Score:      951.00      Matches:      184
XX      Percent Similarity: 93.24%      Conservative: 9
XX      Best Local Similarity: 88.89%      Mismatches: 14
XX      Query Match:      81.84%      Indels:      0
XX      DB:      22      Gaps:      0
XX
XX      US-09-245-198A-2 (1-225) x AAS03964 (1-898)
OY      19 SerleuglySerTPAlarhLreuseralagIngluproSerclngluLeuthrala 38
OY      ||||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      250 AGTTTGGGAGACCGGCGATCGTGTCCGCCAGAGACCTTCCACAGAGAGAGCTGTGCGCA 309
OY      39 GLUAspaAgaAygGluPProPGroGluLeuAsrPrGlnThrClngluSersclnAspAylal 58
OY      ||||| :||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      310 GAGGAGGCGACGAGCCGCTGGAATGTAAATCCCAAGACAGAAAGAACCCAGGATCTCTGCG 369

```

Oy	59	ProheleuengluGluLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArg	78		
Db	370	CGTTTCCCTGAAACCGACTAGTTCGGCCCTCGAGAAAGTCACTTAAAGGCGGAAAAACACGG	429		
Oy	79	ProArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAla	98		
Db	430	GCTGCAAGAGAGATCCGACGCCCATTTATGAAGTTCATCCACGACTCGAGACAGGACGGAGCG	489		
Oy	99	GlnAlaGlyValAspGlyThrValSerGlyTyrGlnGluThrLysIleAsnSerSerSer	118		
Db	490	CAGGACAGGTGGGACGGGACAGTGAAGTGGCTGGAGAAAGCCAGAAATCAACAGCTCCAGC	549		
Oy	119	ProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyr	138		
Db	550	CCCTGGGCGCTACAACCCCGAGATCGGGAGATTATATGTCACCCCGGGCTGGCTCTACTAC	609		
Oy	139	LeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu	158		
Db	610	CTGACCTGTCAAGGTCCTTTGATGAGGGGAAAGGCTGTCTACTCTGAAGCTGCATTTGCTG	669		
Oy	159	ValAsnGlyValLeuAlaLeuAlaArgCysLeuGluGlnPheSerAlaThrAlaIleSerSer	178		
Db	670	GTGAGATGTTGTGCTGGCCCTGCGCTGCTGAGGAATTTCTCAGCCACTGCGGCCAGTTCC	729		
Oy	179	ProGlyProGluLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySer	198		
Db	730	CTCGGGCCCGCCAGCTCCGCTCTCCAGAGTGTCTGGGCTTGGCCCTGGCGCCAGGGTCC	789		
Oy	199	SerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyr	218		
Db	790	TCCCTGGGAGATCCGACCTCCCTCCCTGGGACCATCTCAAGGCTGCCCCCTCTCTCACTTAC	849		
Oy	219	PheGlyLeuPheGlnValHis	225		
Db	850	TTCCGACCTCTCCAGGTTTAC	870		
RESULT 11					
AAAX23424					
ID	AAAX23424	standard; DNA: 1030 BP.			
AC	AAAX23424:				
XX	XX				
DT	18-JUN-1999	(first entry)			
XX	XX				
DE	Human TNRL3 DNA.				
XX	XX				
KW	Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;				
KW	developmental abnormality; gestational abnormality; prostate cancer;				
KW	AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;				
KW	cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;				
KW	apoptosis; human; ss.				
XX	XX				
OS	Homo sapiens.				
XX	XX				
FT	Key	Location/Qualifiers			
FT	CDS	1..627			
FT		/*tag= a			
FT		/product= "TNRL3"			
XX	XX				
PN	MO9911791-A2.				
XX	XX				
PD	11-MAR-1999.				
XX	XX				
PE	04-SEP-1998;	98MO-US18393.			
XX	XX				
PR	05-SEP-1997;	97US-0924634.			
XX	XX				
PA	(UNIM ) UNIV WASHINGTON.				
XX	XX				
PI	Chauchary PM.				
XX	XX				
XX	XX				
WR	WPI: 1999-205191/17.				
OR	P-PSDB: AAW93590.				
OR					

XX New Tumor Necrosis Factor family receptor polypeptides and ligands -  
PT useful for diagnosis and treatment of prostate cancer and  
PT developmental or gestational abnormalities  
XX  
PS Example VII: Fig 13A; 156pp; English.

CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
CC fragments, and isolated TNF related ligands 1 and 3 (TNFR1 and TNFR3) or  
CC their active fragments. APO4 is useful for diagnosing prostate cancer  
CC by determining levels of APO4 in an individual. Prostate cancer can also  
CC be treated using APO4 selective binding agents linked to a therapeutic  
CC moiety. APO4 polypeptides are also useful for identifying selective  
CC binding agents, useful in diagnosis/treatment of disease by binding of  
CC agents to the polypeptide/active fragment which is extracellular, or  
CC expressed on the cell surface. The binding is preferably performed *in*  
CC *vitro*. APO4 polypeptides/active fragments are also useful for screening  
CC for agonists and antagonists by binding and observing the changer in APO  
CC activity. Effective pharmacological agents useful in diagnosis or  
CC treatment of disease are also identified using APO4 polypeptides/active  
CC fragments and APO4 signal transducer molecules that specifically interact  
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO  
CC activity. The method is performed *in vivo* or *in vitro*. APO polypeptides  
CC are all useful as immunogens for preparing antibodies. APO4 is also  
CC useful for diagnosis/treatment of developmental or gestational  
CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
CC MCF-7, and induced apoptosis.

Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Alignment Scores:	
Pred. No.:	1,26e-77
Score:	951.00
Percent Similarity:	92.79%
Best Local Similarity:	88.94%
Query Match:	81.84%
DB:	20
Length:	103
Matches:	185
Conservative:	8
Mismatches:	15
Indels:	0
Gaps:	0

US-09-245-198A-2 (1-225) X AAX23424 (1-1030)

Oy	18	ValSerLeuIglySerTTPAlaIThrLeuSerLacInIuProSerInIuIuLeuThr	37
Db	1	GTCAAGTTTGGGAGACCGGGCATTCGCTGTCCGCCAGGAGCCTGGCCAGAGAGACTGGTG	60
Oy	38	AlaGluAspArgArgGluuProProGluLeuAsnProGlnThrGluGluSerGlnAspVal	57
Db	61	GCAGAGGAGGACACGAGACCCCTGGGAAGTGAATCCCGACAGAGAAACCCAGAGTCTT	120
Oy	58	ValProPheLeuGluGlnLeuValArgProArgArgSerAlaProIysGlyArgGlyAla	77
Db	121	GCGCCTTTCGTAACCCGACACTGATGGCTGTCCGAAAGTCACTTAAAGCCGGAAAAACA	180
Oy	78	ArgProArgArgAlaIleAlaAlaHisIstyGluValHisProArgProGluGlnAspGly	97
Db	181	CGGCGTCGAAGAAGCATGCCATTCATGAATTCATCCAGCACTGGACAGAGACGGA	240
Oy	98	AlaGlnAlaGlyValAspGlyThrValSerGlyTTPcIuGluThrIlysIleasnSerSer	117
Db	241	CGCGAGCGAGGTGTGGACGGGACAGTAGTGCTGGGAGAGAACCAATCAACAGCTCC	300
Oy	118	SerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyr	137
Db	301	AGCCCTTCGCCCTCAACCGGCAATGGGGAGTTTATGACACCCGGCTGGGCTCTAC	360
Oy	138	TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu	157
Db	361	TACCGTAGCTGCAGGTGGCAATTTGATGAGGGGAAGGCTGTCAACCGAAGCTGGACTTG	420
Oy	158	LeuValAsnGlyValIleuAlaLeuAlaArgCysLeuGluIuPheSerAlaThrAlaIleSer	177
Db	421	CTGGGTGATGGTGTCTGGCCCTTCGCTGTGGAGGAATTTTCAGGCACCTGGGGCAGT	480

Qy	178	SerProGlyProGlnLeuATyGleuCYSGlnValSerGlyLeuLeuProLeuArProGly	199
Db	481		540
Qy	198	SerSerLeuArgIleArgThrLeuProTyrPalaHisLeuValAspAlaProPheLeuThr	217
Db	541		600
Qy	218	TyrPheGlyLeuPheGlnValHis	225
Db	601		624
RESULT	12		
ID	ABK11680		
XX	ABK11680	standard; DNA; 2148 BP.	
AC	ABK11680;		
XX	05-JUN-2002	(first entry)	
DE	DNA	encoding tumour necrosis factor variant 2 (TNFv2).	
XX			
KM	Tumour necrosis factor; TNF; pesticide; crop-damaging insect;		
KM	farm animal insect; epithelial morphogenesis; cell-matrix adhesion;		
KM	x-linked anhidrotic (hypohidrotic) ectodermal dysplasia;		
KM	x-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder		
KM	spars hair; sweat gland aberration; endotoxic shock; inflammation;		
KM	hemorrhagic necrosis of tumour; cytotoxicity; TNFv2;		
KM	obesity-linked insulin resistance; gene; ds.		
XX			
OS	Drosophila melanogaster.		
XX			
FH	Key	Location/Qualifiers	
FT	sig_peptide	634..789	
FT		/+tag= a	
FT	CDS	634..1863	
FT		/+tag= b	
FT		/product= "TNFv1"	
FT	mat_peptide	/note= "Tumour necrosis factor variant 1"	
FT		790..1860	
FT		/+tag= c	
FT		/label= mature_TNFv1	
FT	misc_difference	634..1860	
FT		/+tag= d	
FT		/note= "Specifically claimed in claim 22"	
XX			
PN	US2002012968-A1.		
XX			
PD	31-JAN-2002.		
XX			
PE	20-MAR-2001; 2001US-0813329.		
XX			
PR	21-MAR-2000; 2000US-190816P.		
XX			
PA	(CARR/) CARROLL P M.		
PA	(CHEN/) CHEN J		
PA	(RAMA/) RAMANATHAN C S.		
PA	(XIAO/) XIAO H.		
PA	(GUAN/) GUAN B.		
PA	(BOWE/) BOWEN M A.		
XX			
PI	Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA.		
XX			
DR	WPI; 2002-195121/25.		
XX	P-PSDB; AA077718.		
PT	New Drosophila tumour necrosis factor molecule, useful in controlling		
PT	agriculturally important pests, e.g. comprises modifying the growth,		
PT	feeding or reproduction of crop-damaging insects or insects of farm		
XX	animals		
XX			
PS	Claim 2; Fig 3A-C; 119p; English.		

	CC	The invention describes an isolated tumour necrosis factor polypeptide (TNF). The polypeptide and polynucleotide are useful in controlling agriculturally important pests, particularly by modifying the growth, feeding and/or reproduction of crop-damaging insects or insects of farm animals. The polypeptide and polynucleotide are useful for modulating epithelial morphogenesis, cell-matrix adhesion in flies and mammals.
	CC	This the polypeptide and polynucleotide may be useful for treating CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorders, e.g. sparse hair, abnormal or missing teeth or sweat gland aberrations in animals (e.g. insects and potentially humans).
	CC	endoxic shock, inflammation, haemorrhagic necrosis of tumours, cytotoxicity and obesity-linked insulin resistance, all of which involve TNF molecules. This sequence encodes the drosophila melanogaster tumour necrosis factor variant 2 (TNFv2) protein, described in the invention.
SQ	XX	Sequence 2148 BP; 645 A; 510 C; 506 G; 487 T; 0 other;
		Alignment Scores:
	Pred. No.:	0.962                  Length:                  2148
	Score:	113.50                 Matches:                 50
	Percent Similarity:	41.54%                Conservative:           31
	Best Local Similarity:	25.64%                Mismatches:            78
	Query Match:	9.77%                  Indels:                 36
	DB:	24                      Gaps:                   9
US-09-245-198A-2 (1-225)	X	ABK11680 (1-2148)
OY	41	ATGArgGLuPProPGluLeuAsnProGlnThrGluSergIlnAspValValPro-PH 60 :::      :::
Db	1353	CMAAGAGGCCCTTCACCACCTTCAACA-----CGGTGCAGATGGATTCCGCCCA 1403 :::      :::
OY	60	eLeuGLuSLIneuValArGrProAtGrASerSerAlaProLySGlyArGysLaIarGrProAr 80             
Db	1404	TGCGCACCTCCTTAGTCGCAAAAGCCAGATCC-----GAGCAGCTCGAGGCCA-- 1449             
OY	80	gARgALalleaIAaIAahISTyrgIuVaIHISPrOArGrProGIyLInaspLylaaGlnal 100             
Db	1450	-----CGACCCCATTTCCACTTGAGCAGCAGCGCGCGTCACCAAGAAAGTATG-- 1497             
OY	100	aGLyAlaspoljlythrValIseryLYTPRgluGlutThrLySlleasnserseerProle 120         :         :
Db	1498	-GGCTACCCATTAAGAATATGATACATAGAAATGATAACGAGAAGAACTCTTATCAGGGA-- 1554         :         :
OY	120	uARGTyrSPArGrGlnllegIGluPhethrValIlleARGalagIyeuTYrrTyLeuTy 140 :::           :::
Db	1555	-CACTTTCAAACGGCGCATGCCGTCTTGACGGTGCATCCAATACAGGCTTATATACGTATA 1613             
OY	140	rCYsgLIInValIHISpheaspGIugLyLSalValTYrLeuLySLneuspLeuValas 160         :         :
Db	1614	CGCCCAAGATATGCTACAAACAACACTCGCACGAC-----CAGAA 1649         :         :
OY	160	nGIValleuVala-----LeuARgCYsLeuGLuGLuPHeseral 173         :         :
Db	1650	CGGATTATTCCTCTTTCAGAGAGACACTCCATTCCTCTGCACTGCTTGAC----- 1698         :         :
OY	173	aThrAlaIASerSerProGIyProGlnLeuARGleucYsgIInValIseryLyLeuEUPr 193         :         :
Db	1699	-ACGTGCCCCAACCAATGCCCACATTAAGGTGCACACCTCGCACAGCTGTGATGCCA 1757         :         :
OY	193	oleuARgrproILyserSerleuATgIIearyThleu-----ProTrpPLAHISle 210         :         :
Db	1758	CCTGGACGAACGAAAGAGAGATCCATCTGAAGAGACATTCCACAGCATCGCAATGCATTCT 1817         :         :
OY	210	uLYsAlaIalProPheLeuThrTYrrPheGlyLeuPheGlnVal 224         :         :
Db	1818	GCGGAGGAGAAACAACGAGACACTCTTGCGATCTTCAGGTG 1860         :         :
RESULT 13		
ABK11679	ID	ABK11679 standard; DNA; 1221 BP.
XX	ABK11679;	

XX 05-JUN-2002 (first entry)  
 DT DNA encoding tumour necrosis factor variant 1 (TNFV1).  
 DE  
 XX  
 XX Tumour necrosis factor; TNF; pesticide; crop-damaging insect;  
 KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;  
 KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;  
 KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;  
 KW sparse hair; sweat gland aberration; endotoxin shock; inflammation;  
 KW haemorrhagic necrosis of tumour; cytotoxicity; TNFV1;  
 KW obesity-linked insulin resistance; gene; ds.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 FH Key Location/Qualifiers  
 FT sig\_peptide 1..156  
 FT CDS /\*tag= a  
 FT 1..1221  
 FT /\*tag= b  
 FT /product= "TNFV1"  
 FT /note= "Tumour necrosis factor variant 1"  
 FT mat\_peptide 157..1218  
 FT /\*tag= c  
 FT /label= mature\_TNFV1  
 FT misc\_difference 1..1218  
 FT /\*tag= d  
 FT /note= "Specifically claimed in claim 18"  
 XX  
 XX US2002012968-A1.  
 XX  
 XX 31-JAN-2002.  
 PD  
 XX  
 XX 20-MAR-2001; 2001US-0813329.  
 PF  
 XX  
 XX 21-MAR-2000; 2000US-190816P.  
 PR  
 XX  
 XX (CARR/) CARROLL P M.  
 PA (CHEN/) CHEN J.  
 PA (RAMA/) RAMANATHAN C S.  
 PA (XIAO/) XIAO H.  
 PA (GUAN/) GUAN B.  
 PA (BOWEN/) BOWEN M A.  
 XX  
 XX Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;  
 PI  
 XX  
 XX WPI; 2002-195121/25.  
 DR  
 XX  
 XX New Drosophila tumour necrosis factor molecule, useful in controlling  
 PT agriculturally important pests, e.g. comprises modifying the growth,  
 PT feeding or reproduction of crop-damaging insects or insects of farm  
 PT animals  
 XX  
 XX Claim 2; Fig 2A-B; 119pp; English.  
 PS  
 XX  
 XX The invention describes an isolated tumour necrosis factor polypeptide  
 CC (TNF). The polypeptide and polynucleotide are useful in controlling  
 CC agriculturally important pests, particularly by modifying the growth,  
 CC feeding and/or reproduction of crop-damaging insects or insects of farm  
 CC animals. The polypeptide and polynucleotide are useful for modulating  
 CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals.  
 CC Thus the polypeptide and polynucleotide may be useful for treating,  
 CC ameliorating or preventing x-linked anhidrotic (hypohidrotic) ectodermal  
 CC dysplasia-like disorders, e.g. sparse hair, abnormal or missing teeth or  
 CC sweat gland aberrations in animals (e.g. insects and potentially humans),  
 CC endotoxin shock, inflammation, haemorrhagic necrosis of tumours,  
 CC cytotoxicity and obesity-linked insulin resistance, all of which involve  
 CC TNF molecules. This sequence encodes the drosophila melanogaster tumour  
 CC necrosis factor variant 1 (TNFV1) protein, described in the invention.  
 XX  
 XX Sequence 1221 BP; 349 A; 323 C; 312 G; 237 T; 0 other;

Alignment Scores:  
 Pred. No.: 0.792 Length: 1221  
 Score: 111.00 Matches: 44  
 Percent Similarity: 39.81% Conservative: 38  
 Best Local Similarity: 21.36% Mismatches: 84  
 Query Match: 9.55% Indels: 40  
 DB: 24 Gaps: 8  
 US-09-245-198a-2 (1-225) x ABL11679 (1-1221)  
 QY 29 GINGLUPROSERGNGINGLUGLULETHRALAGLUAASPARGARGLUPROPROGLULEUASN 48  
 DB 691 CAGGAAAAGTCCATCCAAATCCAGCACTTCCAAAGAGAGATTCATCCCGCCATCCGCG 750  
 QY 49 PROGLINTHRLUGLUSERGINSAPVALIPROPHLEUGLUGLNULEUVALARGPROARG 68  
 DB 751 CTCTACTCCGCAAGGTGAATCTCTT----- 780  
 QY 69 ARGSERLAPROLYSGLYARGLYSALAARGPROARGALAILAIAHISTYRGLU 88  
 DB 781 -----TCAGCCAGATCCGAGACTCGAGGCCA-----GCAGCCCATTTCCAC 822  
 QY 89 VALHISPROARGPROGLYGLINASPGLYALAGLALAGLYVALASPGLYTHRVALSERGLY 108  
 DB 823 TTGACGACGACGGCGGCGTCCACCAAGAGATGCG--GCTTACCATGGAGATATGATACATA 879  
 QY 109 TRPGLUGLUTHRILYSILEASNSERSESRPROLEUARGTYRSPARGINLEGLYLU 128  
 DB 880 GGAATGATATACGAGCAAACTCTATACAGGA--GCTTCAACGCCGATGGCGTC 936  
 QY 129 PHETHRVALILEARGALAGLYLEUTYRLEUTYRCYSLINVALHISPEASPGLYLU 148  
 DB 937 TTGACGGTGACCAATACAGGCTATATTACGTATACGCCCATATCTACCAACACTCG 996  
 QY 149 LYSALVALTYRLEUYSLEUASPLEUVALASNGLYVALLEUALA----- 164  
 DB 997 CACGAC-----CAGACGGAATTTATCTGCTTTCAAGACGAC 1032  
 QY 165 -----LEUARGCYSLUGLUGLUPHESERLATHRALAASERSESRPROGLYPRO 181  
 DB 1033 ACTCCATTCTCGACGCTCTTGAC-----ACGGTGGCCACCAACATGCCCAT 1080  
 QY 182 GLNLEUARGCYSGLINVALSERGLYLEUPEUPROLEUARGPROGLYSERSESRLEUARG 201  
 DB 1081 AAGGTGCACACCTGCCACAGCAGAGTGTGTGATCCCTCGAAGCAACGAGAGATCCAT 1140  
 QY 202 ILEARGTHREU-----PROTPRALHISLEUYSALALAPROPHLEUPTHRYR 218  
 DB 1141 CTGAAGGACATTCACAAAGATTCGCATGCTGTGGGAGGGAACCAACCGAAGCTAC 1200  
 QY 219 PHEGLYLEUPHEGLINVAL 224  
 DB 1201 TTTGGCATCTTCAAGGTG 1218  
 RESULT 14  
 ABL21473  
 ID ABL21473 standard; DNA; 978 BP.  
 XX  
 AC ABL21473;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 15892.  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ds.  
 OS  
 XX Drosophila melanogaster.  
 PN WO200171042-A2.  
 XX  
 XX 27-SEP-2001.  
 XX

PF 23-MAR-2001; 2001WO-US09231.  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI: 2001-656860/75.  
 DR  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Claim 1: SEQ ID NO 15892; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB101840-AB16175), expressed DNA  
 CC sequences (AB57737-AB872072).  
 CC (AB57737-AB872072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SQ Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.663 Length: 978  
 Score: 110.50 Matches: 49  
 Percent Similarity: 41.62% Conservative: 33  
 Best Local Similarity: 24.87% Mismatches: 81  
 Query Match: 9.51% Indels: 34  
 DB: 23 Gaps: 9  
 US-09-245-198a-2 (1-225) x AB121473 (1-978)  
 OY 41 Argatggluprprogluleuansnproglnthrglulserglaspvalpro-ph 60  
 DB 450 CAAAGAGAGCCCTGCACCACTTCACCA-----CCGTCCAGAAATCATTCGCCCA 500  
 OY 60 eleuugluleuvalargproargargser-----AaProlysglyArglyAlaLar 78  
 DB 501 TCGCCACCTCTAGTCCGCGAAAGTGAAATCTCTTCTTCAGCCAGATCCGAGACTCGAG 560  
 OY 78 gproargatgalaialaialaiahistyrclvalahisproargprogllyglaspqlval 98  
 DB 561 GCCA-----GCCAGCCATTTCACCTTGACAGCAGCGCGCGCACCAAGGAG 608  
 OY 98 aglnaiaaglyvalaspglythvalsercltyrtpgluglunthrlhlyleasnserserse 118  
 DB 609 TATG---GGCTACCATGCGATATGATAGAAATGATATACAGAGAAACTTATCA 665  
 OY 118 rProleuargtyrargprargglnileglyclunthervhvalilearglaiglyleutyty 138  
 DB 666 GGGA---CACTTTCAAACCGCGGATGCGCTCTTGACGGGACCAATACAGGCTATATTA 722  
 OY 138 rleutyrcysglnvalahisphaspqluglylsalvaltyrleuylsleuaspleule 158  
 DB 723 CGTATACGCCCGCATATGCTACACACACTCCGACGAC----- 759  
 OY 158 uvalasnglyvalleuala-----leuarGysleuglucluph 171  
 DB 760 -CAGAGACGATTATTCGCTTTCAGAGAGACACATTCCTCGAGTGGTGAAC----- 813  
 OY 171 eserlarhralaialaserserprogllyproglneuarGysglnvalserglyle 191  
 DB 814 -----ACGGTGGCCCGACCATGCGCATTAAGGTGACACACTCGCCACAGAGTGTCT 866

OY 191 ubeuProleuargprogllySerleuarglileargthleu-----ProTrpAl 208  
 DB 867 GATCCACCTGAGACGAAACGAGAGATTCATGGAAGGACATTCACACGATCGCAATGC 926  
 OY 208 ahisleuysalaalaprophleuThrTyRPhcglyleupheglnval 224  
 DB 927 AGTCTCGCGGAGGAGCAACCAAGCTACTTGGCATCTTCAGATG 975  
 RESULT 15  
 AA41377  
 ID AA41377 standard; cDNA; 1630 BP.  
 AC AA41377;  
 XX  
 DT 08-OCT-1998 (first entry)  
 XX  
 DE NF-kB receptor activator RANK ligand (RANKL) encoding cDNA.  
 XX  
 KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;  
 KW immune response; inflammatory response; toxic shock; sepsis;  
 KW RANKL; RANK ligand; tumour necrosis factor; TNF; mouse; ss.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 3..887 /tag= a  
 FT /product= "murine RANKL (ligand for RANK)"  
 XX  
 PN W09828426-A2.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-US23775.  
 XX  
 PR 14-OCT-1997; 97US-0064671.  
 PR 23-DEC-1996; 96US-005978.  
 PR 07-MAR-1997; 97US-0813509.  
 XX  
 PA (IMMUNEX CORP.  
 XX  
 PI Anderson DM, Galibert LJ, Maraskovsky E;  
 XX  
 DR WPI: 1998-377657/32.  
 DR P-PSDB; AAM69956.  
 XX  
 PT New isolated ligand for receptor activator of NF-kappa B - used to  
 PT develop products for augmenting an immune response for inhibiting an  
 PT inflammatory response and for protection of cells  
 XX  
 PS Claim 25; Pages 55-57; 80pp; English.  
 XX  
 CC This cDNA encodes a murine RANKL, a ligand for the RANK (receptor  
 CC activator of necrosis factor-kappa B (NF-kB)) polypeptide. RANK is a  
 CC member of the tumour necrosis factor (TNF) family. A soluble RANK  
 CC may be used for inhibiting activation of NF-kB, by contacting a cell  
 CC expressing membrane-associated RANK with a soluble RANK which binds to  
 CC RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be  
 CC used to induce maturation of dendritic cells and enhance their  
 CC allo-stimulatory capacity, thereby augmenting an immune response. The  
 CC soluble RANK polypeptide composition may also be used for regulating an  
 CC immune or inflammatory response. Inhibition of NF-kB by RANK antagonists  
 CC may be useful in ameliorating negative effects of an inflammatory  
 CC response that result from triggering of RANK, e.g. in treating toxic  
 CC shock or sepsis, graft-versus-host reactions, or acute inflammatory  
 CC reactions. They can also be used in adjunct therapy for disease  
 CC characterised by neoplastic cells that express RANK. RANKL polypeptides  
 CC can also be used to identify inhibitors of RANK and thus inhibitors of  
 CC an inflammatory response, and also for protecting RANK-expressing cells  
 CC from the negative effects of chemotherapy or the presence of high levels  
 CC of TNF-alpha. The products can also be used for detection and drug  
 CC screening.  
 XX

Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other;

## Alignment Scores:

Pred. No.:	2,95	Length:	1630
Score:	106.50	Matches:	67
Percent Similarity:	38.04%	Conservative:	38
Best Local Similarity:	24.28%	Mismatches:	106
Query Match:	9.17%	Indels:	65
DB:	19	Gaps:	13

US-09-245-198a-2 (1-225) x AAV41377 (1-1630)

```

QY 3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGlySer 22
    |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 78 TCCATGTTCTGGCCCTCGGGGCTGGGACTGGCCAGGCTGGCTGCAGCATCGCTCG 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 23 TrpAlaThrLeuSerAlaGln---GluProSerGlnGluGluLeuThrAlaGluAspArg 41
    ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 138 TTCCTGACTTTCGAGCGGAGATGATCCTAACAGA-----ATATCAGAAAGACAGC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 42 -----ArgGluProGluLeuAsnProGlnThr 51
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 ACTCAGCTCTTTATAGATCTCGAGACTCCATGAACGAGATTGTCAGAGCTGACT 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 52 GluGluSerGlnAspValPro----- 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 CTGGAGAGTGACAGACACTACTCTCTCGAGAGAGATGAACACGCTTTCAGGGG 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 -----PheLeuGlnGlnLeuValArgProArgArg-----SerAlaProLys 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 309 GCCGTGCAAGAGAACTGCAACACATTGTGGGCCACAGGCTTCTCAGAGAGCTCCAGCT 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 74 -----GlyArgLysAlaArgProArgArgAlaIleAla 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 369 ATGATGGAAGCTCATGTTGATGTGGCCACGAGCAAGCCTGAGGCCACGCTTT 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 85 AlaHisTrpGluValHisProArgProGlyGlnAspGlyValGlnAlaGlyValAspGly 104
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 429 GCACACCTCATCAATCAATGCTGCCAGCATCCATCGGGTCCCAT-----AAAGTC 479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 ThrValSerGlyTrpGluGlu-----ThrLysIleAsnSerSerSerProLeu 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 480 ACTCTGTCTCTTGTGTAACGATCGAGGCTGGGCCAAGATCTCTAACATGACG----- 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 -----TTAAGCAACGAAACAACTAAGGTTAACCAAGATGGCTTCTATTACCTGTAC 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 141 -----CysGlnValHisPheAspGlyLysAlaVal-----TyrLeuLys 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 585 GCCAATCTTGTCTTCCGATCATGAACATCGGAAGCGTACCTACAGACTATCTTCAG 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 LeuAspLeuLeuVal-----AsnGlyValLeuAlaLeuArgCysLeuGluGlu 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 645 CTGATGGTGTATGTCGTTAAACCACGATCAAAATCCCAAGTCTCATTAACCTGATGAAA 704
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 PheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeu-----CysGlnVal 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 705 GGAGGGAGCAGCAAAACTGTCGGGCAATTCGAAATTCACCTTTTATCCATAAATGTT 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 SerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTrpAla 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 765 GGGGGAATTTTCAAGCTCGAGCTGGAGAAATTTGCAATTCAGAGTGTCCAAACCCCTTCC 824
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 209 HisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 825 CTGCTGGATCCGGATCAAGATGCAGCTACTTGGGGCTTTCAAAAGTT 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: March 31, 2003, 02:40:55  
Job time : 241.957 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 00:53:03 ; Search time 2837.48 Seconds

(without alignments)  
2307.731 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLISGLALACGLLVVSL.....PWAHLKAPFLTYFLQVH 225

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Command line parameters:

-MODE=frame+pn.model -DEV=xlh  
-Q=/gen2.1/USPTO/US09245198/runat\_24032003\_163553\_25950/app.query.fasta.1.846  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human4.0.cdi -LIST=45  
-LOCALIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09245198 -ECGN.1.1.3077 -runat\_24032003\_163553\_25950 -NCPU=6 -ICPU=3  
-NO\_XLPTX -NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mnu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hg\_hum:\*  
31: em\_hg\_inv:\*  
32: em\_hg\_other:\*  
33: em\_hg\_mus:\*  
34: em\_hg\_pln:\*  
35: em\_hg\_rod:\*  
36: em\_hg\_mam:\*  
37: em\_hg\_vrt:\*  
38: em\_sy:\*  
39: em\_hgo\_hum:\*  
40: em\_hgo\_mus:\*  
41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162	100.0	1168	10	AF030100	AF030100 Mus muscu
2	1020	87.8	1236	6	AR140407	AR140407 Sequence
3	1020	87.8	1306	9	AF030099	AF030099 Homo sapi
4	1020	87.8	1353	6	AX201324	AX201324 Sequence
5	1020	87.8	1368	9	AF055872	AF055872 Homo sapi
6	951	81.8	898	6	AX180714	AX180714 Sequence
7	945.5	81.4	1651	9	BC019047	BC019047 Homo sapi
8	602.5	51.9	203083	2	AC069459	AC069459 Mus muscu
9	602.5	51.9	234182	10	AL603707	AL603707 Mouse DNA
10	549.5	47.3	177703	2	AC016876	AC016876 Homo sapi
11	549.5	47.3	215795	2	AC127470	AC127470 Pan trogl
12	545	46.9	177555	2	AC130192	AC130192 Sus scrof
13	544.5	46.9	161428	2	AC126925	AC126925 Canis fam
14	536.5	46.2	138792	2	AC119115	AC119115 Rattus no
15	503	43.3	153553	2	AC126921	AC126921 Bos tauru
16	425.5	36.6	184026	2	AC098923	AC098923 Rattus no
17	207	17.8	203281	2	AC126237	AC126237 Canis fam
18	113.5	9.8	1656	3	AB073865	AB073865 Drosophi
19	110.5	9.5	1221	3	AY119233	AY119233 Drosophi
20	110.5	9.5	2159	3	AF521176	AF521176 Drosophi
21	109.5	9.4	892	5	GGA24345	AJ243435 Gallus ga
22	109.5	9.4	108967	9	AL353138	AL353138 Human DNA
23	108.5	9.3	42210	1	SC1C2	AL031124 Streptomy
24	107.5	9.3	13243	1	AE004602	AE004602 Pseudomon
25	107	9.2	178262	2	AC061974	AC061974 Homo sapi
26	107	9.2	179383	2	AC060789	AC060789 Homo sapi
27	106.5	9.2	1630	6	AR156433	AR156433 Sequence
28	106.5	9.2	1630	6	AR164147	AR164147 Sequence
29	106.5	9.2	1630	6	AX147987	AX147987 Sequence
30	106.5	9.2	1694	6	AX451897	AX451897 Sequence
31	106.5	9.2	2225	10	AF019048	AF019048 Mus muscu
32	106.5	9.2	4034	6	AX200995	AX200995 Sequence
33	106.5	9.2	4034	6	AX267730	AX267730 Sequence
34	106.5	9.2	4894	6	AX200993	AX200993 Sequence
35	106.5	9.2	4894	6	AX267728	AX267728 Sequence
36	106.5	9.2	6976	6	AX200996	AX200996 Sequence
37	106.5	9.2	6976	6	AX267731	AX267731 Sequence
38	106.5	9.2	157988	2	AC096533	AC096533 Homo sapi
39	106	9.1	177485	2	AC090975	AC090975 Homo sapi
40	106	9.1	207585	2	AC073779	AC073779 Mus muscu
41	106	9.1	222037	2	AC073801	AC073801 Mus muscu
42	105.5	9.1	4412	9	AY070219	AY070219 Homo sapi
43	105.5	9.1	6305	9	HMU1YTOXBR	L11016 Homo sapien
44	105.5	9.1	24526	6	AX472617	AX472617 Sequence
45	105.5	9.1	62370	2	AL451008	AL451008 Homo sapi

RESULT 1

#### ALIGNMENTS

AF030100  
LOCUS AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
DEFINITION Mus musculus TWEAK mRNA, partial cds.  
ACCESSION AF030100  
VERSION AF030100.1 GI:2707220  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 1168)  
AUTHORS Chicheportiche, Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H., Hession, C., Garcia, I. and Browning, J.L.  
TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis  
J. Biol. Chem. 272 (51), 32401-32410 (1997)  
JOURNAL J. Biol. Chem. 272 (51), 32401-32410 (1997)  
MEDLINE 98070415  
PUBMED 9405449  
REFERENCE 2 (bases 1 to 1168)  
AUTHORS Chicheportiche, Y., Blixler, S., Tizard, R. and Browning, J.  
TITLE Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA  
JOURNAL  
FEATURES  
source location/Qualifiers  
1. 1168  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/cell\_type="peritoneal macrophages"  
/note="ligand in the TNF family; secreted protein"  
/codon\_start=2  
/product="TWEAK"  
/protein\_id="AAC53517.1"  
/db\_xref="GI:2707221"  
/translation="VLSIGLQALLVYVSGSMATLSAQPSQELTAERREP  
PELNPOTBESQDVVPELEQVPRRSAPKRRKRRPRALAHAEVHPRPDQAQCY  
DGVNSWMEETKINSSSPLEARDQIGETVIRALVLYLCVHDEKAVYLKLDLVN  
GVLALRCLEFEFATAASSPQPLRLCOVSGILPLRGSSLSIRTLPMAHKAPFLTY  
FGLEOVH"  
BASE COUNT 242 a 360 c 298 g 268 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,47e-79 Length: 1168  
Score: 1162.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 10  
US-09-245-198a-2 (1-225) x AF030100 (1-1168)  
Qy 1 Valueusertleuglyleualaleualacysleuglyleuleuleuvalvaliserleu 20  
Db 2 GTGCTGAGCCCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 61  
Qy 21 Glysertrpalathrlseuseralaglnluproserlnglnlupserlnglnlupser 40  
Db 62 GGGAGTGGGCAAGCGTGTGCCAGAGACCTTCAGAGAGAGTGCAGAGAGAGAC 121  
Qy 41 ArgArgcluproproglyleuuanpproglnthrglnlupserglnaspvalvalprophe 60  
Db 122 CGCGGGAGAGCCCTGAACTGAATCCAGACAGAGAAAGCCAGATGTGGTACCTTTC 181  
Qy 61 leuglnlupserleuvalargproargargseralargprolysglyargylsalargproarg 80  
Db 182 TTGGAAACACTAGTCCGGCTCGAAGAGAGTCTCTAAAGCCGGAAGCGCGGCTTCGC 241  
Qy 81 Argglatleualahstlyrglyvalahisproargproglgynaspvalvalaglnala 100  
Db 242 CGAGCTTTGGCAGCCCTTATGAGCTTCACTCTGGCCAGAGAGATGGAGCACAAGCA 301  
Qy 101 glyvalaspgllythrvaliserglytrpdluglnluphrlyslleasenserserproleu 120

|||||  
Db 302 GGTGTGATGGAGACAGTACGTGGCTGGAGAGACCAAAATCAACAGCTCCAGCCCTCTG 361  
Qy 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140  
Db 362 CGTACGACCGCCAGATGGGGAATTTACAGTCATCGAGGCTGGCTCTACTACTGTATC 421  
Qy 141 CysGlnValHisPheAspGluGlyAlaValTyrLeuLysLeuAspLeuValAsn 160  
Db 422 TGTCAAGTGCACCTTGTATGAGGAAAGCGTCTACTGAAGCTGGAGCTTGGTGTAAC 481  
Qy 161 GlyValleuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaAspSerProGly 180  
Db 482 GGTGTGCTGGCCCTGGCTGGCTGGAGAAATTCACACCACAGACGACACTCTCCCTGG 541  
Qy 181 ProGlnLeuArgLeuGlyGlnValiserGlyleuleuLeuProleuArgProGlySerSerleu 200  
Db 542 CCCAGCTCCGTTGTGTGCCAGGTGTCTGGCTGTGGCTGGCCAGAGGCTTCCCTT 601  
Qy 201 ArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGly 220  
Db 602 CGGATCGGCAACCTCCCTGGGCTCATCTTAAGGCTGCCCTTCTTAACCTACTTGA 661  
Qy 221 leupheglnvalhis 225  
Db 662 CTCTTCAAGTTCAC 676  
RESULT 2  
AR140407  
LOCUS AR140407 1236 bp DNA linear PAT 16-JUN-2001  
DEFINITION Sequence 1 from patent US 6207642.  
ACCESSION AR140407  
VERSION AR140407.1 GI:14482903  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1236)  
AUTHORS Willey, S.R.  
TITLE Member of the TNF family useful for treatment and diagnosis of disease  
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;  
FEATURES  
source location/Qualifiers  
1. 1236  
/organism="unknown"  
BASE COUNT 225 a 416 c 358 g 237 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.38e-69 Length: 1236  
Score: 1020.00 Matches: 199  
Percent Similarity: 92.86% Conservative: 9  
Best Local Similarity: 88.84% Mismatches: 16  
Query Match: 87.78% Indels: 0  
Gaps: 0  
DB: 6  
US-09-245-198a-2 (1-225) x AR140407 (1-1236)  
Qy 2 leuserleuglyleualaleualacysleuglyleuleuleuvalvaliserleugly 21  
Db 76 CTGGGCTGGGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 135  
Qy 22 sertrpalathrlseuseralaglnluproserlnglnlupserlnglnlupser 41  
Db 136 AGCGGGCATCGCTGTCCGCCAGAGCTTGCCAGAGAGAGTGTGGCAGAGAGAGAC 195  
Qy 42 ArgglupProGlyleuuanpProglnthrglnlupserglnaspvalvalpropheleu 61  
Db 196 CAGAGCCCGTCGAATCGAATCCAGACAGAGAAAGCAGATCCTGGCCTTCTCTG 255  
Qy 62 GluGlnleuValArgProArgArgSerAlaProLysGlyArgylsalargproargarg 81  
Db 256 AACCGACTAGTTCGCTCGAAGAGTGCACCTTAAGCGCGGAAMACAGCGGCTCGAAGA 315

QY 82 ALaIleAlaIaHISrTgCluValHisProArGProGlyGlnAspGlyAlaGlnagly 101  
|||||  
Db 316 GCGATCGCAGGCCATTATGATGATCATCCAGACCTGGACAGAGGCGAGGT 375  
QY 102 ValAspGlyThrValSerGlyTrpGlnGluThrLysIleAsnSerSerProLeuArg 121  
|||||  
Db 376 GTGGACGGGACAGTAGTGCTGGGAGGAGCAAGATCAACAGCTCCAGCCCTCGGCC 435  
QY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141  
|||||  
Db 436 TCAACACCGGCAGATCGGGAGATTATACACCCGGGCGGCTCTACTACCTGACTGT 495  
QY 142 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
|||||  
Db 496 CAGGTGCACTTGATGATGAGGAGGAGGCTGTCTACTGAACTGACTGCTGTGATGGT 555  
QY 162 ValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaIleSerSerProGlyPro 181  
|||||  
Db 556 GTGCTGGCCCTCGCTGGCTGGAGGAATCTCAGCCACTGGCGGAGTCCCTCGGGCC 615  
QY 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
|||||  
Db 616 CAGCTCCGCTCTGCGCAGGTGTCTGGGCTGTGGCCCTGGCCAGGGTCTCCCTGCGCG 675  
QY 202 IleArgThrLeuProThrAlaHisLeuLysAlaIleProPheLeuThrTyrPheGlyLeu 221  
|||||  
Db 676 ATCCGCACCCCTCCCTGGGCGCATCTCAAGGCTGCCCTCTCCTCACTTCCGAGCTC 735  
QY 222 PheGlnValHis 225  
|||||  
Db 736 TTCGACGTTTCAC 747  
RESULT 3  
AF030099 1306 bp mRNA linear PRI 20-DEC-1997  
LOCUS Homo sapiens TWEAK mRNA, complete cds.  
DEFINITION AF030099  
ACCESSION AF030099.1 GI:2707218  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 1306)  
Hession, C., Garcia, I., and Browning, J.L.  
TWEAK, a new secreted ligand in the tumor necrosis factor family  
that weakly induces apoptosis  
J. Biol. Chem. 272 (51), 32401-32410 (1997)  
MEDLINE 98070415  
PUBMED 9405449  
REFERENCE 2 (bases 1 to 1306)  
Bourdon, P., Hession, C., Tizard, R. and Browning, J.  
Direct Submission  
TITLE Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,  
JOURNAL Cambridge, MA 02142, USA  
FEATURES  
source  
location/Qualifiers  
1. 1306  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17p13"  
/tissue\_type="tonsil"  
18..767  
/note="Ligand in the TNF family; secreted protein; start  
codon not verified experimentally"  
/codon\_start=1  
/product="TWEAK"  
/protein\_id="AAC51923.1"  
/db\_xref="GI:2707219"  
/translation="MAARSSQRRRGRRGPEPTALVPLALGLALACTGLAVSL  
GSRASLSHOEPAOEEIYVAEEDPSSELNPQEEESODPAPFNLRLVPRRSAPKGRKTR

ARRAIAAHYEVPBGODGAGVGTGSCGEARINSSSPLRYNRQIGERTYRACL  
YLYCOVHFDEGKAVLYKLDLVGVVALRCLFEFSATMAASLSLPQRLCOVSLALL  
RFSLSLRTRTPMHLKRAPLTYRGLFQVH"  
BASE COUNT 247 a 434 c 368 g 257 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9,95e-69 Length: 1306  
Score: 1020.00 Matches: 199  
Percent Similarity: 92.86% Conservative: 9  
Best Local Similarity: 88.84% Mismatches: 16  
Query Match: 87.78% Indels: 0  
DB: 9 Gaps: 0  
US-09-245-198a-2 (1-225) x AF030099 (1-1306)  
QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21  
|||  
Db 93 CTGGGCGCTGGGCGCTGGGCGCTGCTGCGCTGCTGCTGCGGCTGCTGCTGCTGCTG 152  
QY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41  
|||  
Db 153 ACCCGGGCATCTGCTGCCGCCAGAGCCTGCCAGAGAGACTGTGGCAGAGAGAGAC 212  
QY 42 ArgGluProProGlnLeuAsnProGlnThrGlnGluSerGlnAspValValProPheLeu 61  
|||||  
Db 213 CAGGACCCGTCGGAACGTGATCCAGACAGAGAGAGAGAGAGAGATCTGCGCTTCCG 272  
QY 62 GlnGlnLeuValArgProArgArgSerAlaPolysGlyArgAlaAlaArgProArg 81  
|||||  
Db 273 AACCGACTAGTTCGGGCTCGCAGAAAGTGCACCTAAAGCGGAGAAACAGCGCTCGAGA 332  
QY 82 ALaIleAlaIaHISrTgCluValHisProArGProGlyGlnAspGlyAlaGlnagly 101  
|||||  
Db 333 GCGATCGCAGGCCATTATGATGATCATCCAGACCTGGACAGAGGCGAGCGAGGT 392  
QY 102 ValAspGlyThrValSerGlyTrpGlnGluThrLysIleAsnSerSerProLeuArg 121  
|||||  
Db 393 GTGGACGGGACAGTAGTGCTGGGAGGAGCAAGCAAGATCAAGCTCCAGCCCTGTGGCG 452  
QY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141  
|||||  
Db 453 TCAACACCGGCAGATCGGGAGATTATGATGATCATCCAGCCCGGCGCTCTACTACCTGATG 512  
QY 142 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
|||||  
Db 513 CAGGTGCACTTTGATGAGGGAGGAGGCTGTCTACTGAACTGACTGTCTGTGATGATGT 572  
QY 162 ValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaIleSerSerProGlyPro 181  
|||||  
Db 573 GTGCTGGCCCTCGCTGCTGGAGGAATCTCAGCCACTGCGGCGCACTTCCCTCGGGCC 632  
QY 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
|||||  
Db 633 CAGCTCGGCGCTGCTGGCAGGTGTCTGGGCTGTGGCCCTGGCGCAGGGTCTCCCTCGCGG 692  
QY 202 IleArgThrLeuProThrAlaHisLeuLysAlaIleProPheLeuThrTyrPheGlyLeu 221  
|||||  
Db 693 ATCCGCACCCCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTCCTCACTTCCGAGCTC 752  
QY 222 PheGlnValHis 225  
|||||  
Db 753 TTCGACGTTTCAC 764  
RESULT 4  
AX201324 1353 bp DNA linear PAT 30-AUG-2001  
LOCUS Sequence 3 from Patent W00153486.  
DEFINITION AX201324  
ACCESSION AX201324  
VERSION AX201324.1 GI:15391154  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE	Eukaryota; Metazoa; Chordata; Ctenaria; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	1 (bases 1 to 1353) Ashkenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pan, D., Pilti, R.M., Roy, M.A., Smith, V., Stone, D.M., Watanabe, C.K. and Wood, W.I.
TITLE	Compositions and methods for the treatment of tumour
JOURNAL	Patent: WO 0153486-A 3 26-JUL-2001; Genentech, Inc. (US)
FEATURES	Location/Qualifiers source 1..1353 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	257 a 443 c 389 g 264 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.03e-68
Score:	1020.00
Percent Similarity:	92.86%
Best Local Similarity:	88.84%
Query Match:	87.78%
DB:	Gaps: 0
US-09-245-198A-2 (1-225) x AX201324 (1-1353)	
QY	2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
Db	133 CTGGGCGCTGGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGG 192
QY	22 SerTrpAlaThrLeuSerAlaGlnGlnPProSerGlnGlnGlnLeuThrAlaGly 41
Db	193 ACGCGGCGCTGGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 252
QY	42 ArgGlnProProGlnLeuAsnProGlnThrGlnGlnSerGlnAspValAlaProPheLeu 61
Db	253 CAGGACCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCT 312
QY	62 GlnGlnLeuValArgProArgArgSerAlaProGlyArgGlyArgGlyAlaArgProArgArg 81
Db	313 AACCGACGAGTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 372
QY	82 AlaIleAlaAlaHisTyrGlnValHisProArgProGlnGlnAspGlyAlaGlnAlaGly 101
Db	373 GCGATCGGACGCCATTATGAACTTATCCACACACTGGACAGAGCGAGCGAGCGAGCT 432
QY	102 ValAspGlyThrValSerGlyTrpGlnGlnThrLysIleAsnSerSerSerProLeuArg 121
Db	433 GTGGACGGGACAGTGAAGCGCTGGGAGGAGCAACGATCAACAGCTCCAGCCCTCGCCG 492
QY	122 TyrAspArgGlnIleGlyGlnPheThrValIleArgAlaGlyLeuThrTyrLeuThrCys 141
Db	493 TACAAACCGCAATGGGGAGTTTATGACACCCGGGCGTCTACTAACCTGTACTCT 552
QY	142 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuLeuValAsnGly 161
Db	553 CAGGTGCACTTGTGATGAGGGAGGAAGCGTCTCACTGAACTGGACTGCTGGTGGATGGT 612
QY	162 ValLeuAlaLeuAlaArgCysLeuGlnGlnPheSerAlaThrAlaAlaSerSerProGlyPro 181
Db	613 GTGGCTGGCGCTGGCGCTGGAGGAAATTCAGCGACAGCGGCGGTTCCCTCGGGGCC 672
QY	182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db	673 CAGCTCCGCGCTGGCAGGCTGTGGGCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGG 732
QY	202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
Db	733 ATCCGACACCTCCCTGGGCGCATCTCAAGGCTGCCCTCTCTCACTCACTTCGAGATC 792
QY	222 PheGlnValHis 225
Db	793 TTCGAGGTTTAC 804

[illegible]

```

Qy 62 GlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
Db 313 AACGACAGTGGCTGGCTGGAGAAAGTGCACCTAAAGCCGGAACACGGGCTGAAAGA 372
Qy 82 AlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 101
Db 373 GCGATCGCAGCCCATTTATGAGTTCATCCAGACCTGGACAGAGGAGCGAGCGAGT 432
Qy 102 ValAspGlyThrValSerGlyTyrProGluLutThrLysIleAsnSerSerProLeuArg 121
Db 433 GTGGCGGGACAGTAGTGCTGGAGGAGCAAGATCAACAGCTCCAGCCCTCTGGCG 492
Qy 122 TyrAspArgGlnIleGlyLutPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
Db 493 TACAACCCCGCAGATGGGGAGTTTATGTCACCCGGGCTGGGCTTACTACTGACTGCT 552
Qy 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
Db 553 CAGGTGCACTTGTGATGAGAGGGGAGGCTGTCTACTGAAAGCTGAGCTGCTGGTGATGT 612
Qy 162 ValIleAlaLeuArgCysLeuGlnGluLutPheSerAlaThrAlaIleAsnSerProGlyPro 181
Db 613 GTGCGCCCTGGCTGGCTGGAGAAATCTCAGCCACTGGCGGAGTTCCCTCGGGCCC 672
Qy 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db 673 CAGCTCGCCCTCTGTGCAGGTGTCTGGGCTGTGGCCCTGGCGGAGGCTCCCTCGCGG 732
Qy 202 IleArgThrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
Db 733 ATCCCGACCTCTCCCTGGGCGCCATCTCAAGGCTGCCCTTCCTCACTTCCGAGCTC 792
Qy 222 PheGlnValHis 225
Db 793 TTCCAGGTTTCAC 804

RESULT 6
AX180714 898 bp DNA Linear PAT 06-AUG-2001
LOCUS AX180714
DEFINITION Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
ORGANISM
SOURCE
REFERENCE 1 (bases 1 to 898)
AUTHORS
Tweak receptor
TITLE
JOURNAL
IMMUNEX CORPORATION (US)
FEATURES
Source
Location/Qualifiers
1..898
/organism="synthetic construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=1
/protein_id="CAC50004.1"
/db_xref="gi:15132571"
/translation="MATGSRPTSLDLAFGLCTLPMLOEGSATSSDRMKOIEDKIEEILS
KIYHENEIARIKILIGERTSSLSRASLSAOBPAOBEIYAEDODSEINPOTEEES
ODPAPFLNRLVRRPRSA PKGKTRARARAI AHAYHVPHPDGDGADGVATGMEEA
RINSSSPLRVROJGEPTVTRAGIYILXCOVHPDEGANVYIKDLIDLGVATLAKLEE
FSATPAASLSLGQLKLCOVSGSLAPSSSLRIRLPMVHMLKAAPFLTYFGLEOVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

```

Alignment Scores: 1.15e-63 Length: 898  
 Score: 951.00 Matches: 184  
 Percent Similarity: 93.24% Conservative: 9

```

Best Local Similarity: 88.89% Mismatches: 14
Query Match: 81.84% Indels: 0
DB: 6 Gaps: 0

US-09-245-198a-2 (1-225) x AX180714 (1-898)
Qy 19 SerLeuGlySerThrAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAla 38
Db 250 AGTTTGGGAGCCGGGCGCATTCGCTGCCCGCCAGAGACCTGCCAGAGAGAGCTGGTGCA 309
Qy 39 GlnAspArgArgGluProProGluLeuAsnProGlnThrGlnGluSerGlnAspVal 58
Db 310 GAGGAGAGCAGAGACCCCGGGAAGTGAATCCCGACAGAAAGAAAGCAGAGATCTCGC 369
Qy 59 ProPheLeuGlnGluLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArg 78
Db 370 CTTTTCGTAACCGACTAGTTCGCTCGCAAGATGCACTTAAAGCCGGAACACAGCG 429
Qy 79 ProArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 98
Db 430 GCTCAAGAGCATGCGCAGCCCATTTATGAGTTCATCCAGACCTGGACAGAGCGAGCG 489
Qy 99 GlnAlaGlyValAspGlyThrValSerGlyTyrProGluLutThrLysIleAsnSerSer 118
Db 490 CAGCGAGGTGTGGACGGAGAGTAGTGGCTGGAGAAAGCCAGATCAACAGCTCCAGC 549
Qy 119 ProLeuArgTyrAspArgGlnIleGlyLutPheThrValIleArgAlaGlyLeuTyrTyr 138
Db 550 CCTTGGCCCTACAAACCGCAGATCGGGAGTTATAGTACACCGGGGCTGGGCTTACTAC 609
Qy 139 LeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 158
Db 610 CTGTACTGTGACAGTGCACCTTTGATGAGGGAAAGCTGTCTACTACAGCTGAGCTGCTG 669
Qy 159 ValAsnGlyValLeuAlaLeuArgCysLeuGlnGluLutPheSerAlaThrAlaIleSer 178
Db 670 GTGATGTGTGTGCTGGCTGGCTGGCTGGAGAAATCTCAGCCACTGGCGGCGAGTTCC 729
Qy 179 ProGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
Db 730 CTGGGCCCCAGCTCGGCTGGCTGGAGGTGTGGGCTGTGGCCCTGGGCGGAGGTC 789
Qy 199 SerLeuArgIleArgThrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyr 218
Db 790 TCCCTGGCGGATCCGACCCCTCCCGGGGCCATCTCAAGGCTGCCCTCTCTCACTAC 849
Qy 219 PheGlyLeuPheGlnValHis 225
Db 850 TTGGGACTCTTCCAGGTTTCAC 870

RESULT 7
BC019047 1651 bp mRNA Linear PRI 11-DEC-2001
LOCUS BC019047
DEFINITION Homo sapiens, similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047
VERSION BC019047.1 GI:17512138
KEYWORDS
ORGANISM
SOURCE
Homo sapiens.
Homo sapiens.
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homini; Homo.
REFERENCE 1 (bases 1 to 1651)
AUTHORS
Straussberg, R.
TITLE
Direct Submission
JOURNAL
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov

```

Tissue Procurement: Louis Staudt  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: [http://www.nisc.nih.gov/nisc\\_mgc@nigr1.nih.gov](http://www.nisc.nih.gov/nisc_mgc@nigr1.nih.gov)

Contact: [nisc\\_mgc@nigr1.nih.gov](mailto:nisc_mgc@nigr1.nih.gov)  
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.M., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dierich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stancil, P., Thomas, P.J.,  
 Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 30 Row: P Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 4507596.

# FEATURES

Location/Qualifiers  
 1..1651  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:20669 IMAGE:4766071"  
 /issue\_type="Primary B-Cells from Tonsils"  
 /clone\_id="NIH.MGC\_48"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 106..510  
 /codon\_start=1  
 /product="Similar to tumor necrosis factor (ligand)  
 superfamily, member 12"  
 /protein\_id="AAH19047.1"  
 /db\_xref="GI:17512139"  
 /translation="MAARRQRRRGRGEPGALLVPLALGSLALACILLLAVSL  
 GSRASLSAQEPAGEELVADDDPSILNPOTESTDPPAPFLNRLVPRRSAPKGRKTR  
 ARRAIAAHVEVPRPGODGADQADGTTCLRP"

CDS  
 BASE COUNT 344 a 517 c 481 g 309 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.76e-63 Length: 1651  
 Score: 945.50 Matches: 198  
 Percent Similarity: 68.09% Conserved: 9  
 Best Local Similarity: 65.13% Mismatches: 17  
 Query Match: 81.37% Indels: 81  
 DB: 9 Gaps: 1

US-09-245-198a-2 (1-225) x BC019047 (1-1651)

QY 2 LeuSerLeuGluLeuAlaLeuAlaCysLeuGluLeuLeuValValSerLeuGly 21  
 DB 181 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 240  
 QY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41  
 DB 241 AGCGGGGCAATCGCTGTCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 QY 42 ArgGluProGluLeuAsnProGlnThrGlnGluSerGlnAspValValProPheLeu 61  
 DB 301 CAGGAGCCGCGGAGACTGATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 QY 62 GluGluLeuValArgProArgArgSerAlaProGlyGlyArgLysAlaArgProArgArg 81  
 DB 361 AACCGACTATGTCGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
 QY 82 AlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 100  
 DB 421 GCGATGCGACCCATATGATGATTCACGACGACGACGACGACGACGACGACGAC 479  
 QY 100 ----- 100

DB 480 TGGAGGTACACAACTTGTCTGAGGCCATGAGATTAAGTGTGGAGCCAAAGATTGAA 539  
 QY 100 ----- 100  
 DB 540 CCCAGTAGAATGTGCTGTGTACTGTGACATGCTGTGCATGAAGGCGAGTGTGCTG 599  
 QY 100 ----- 100  
 DB 600 CAGGGGTGAGGGGTCCATGACGAGGGCCACATCCAAAAGGAGAGAGAAATTCCAGAA 659  
 QY 101 -----Gly 101  
 DB 660 AGAAGAGAGACACATCTCCACCATTAACAGAGGTCAAAGAGAGAACAGACCAAGT 719  
 QY 102 ValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeuArg 121  
 DB 720 GTGGAGGGGACAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 779  
 QY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrrTyrrTyrcys 141  
 DB 780 TACACGGCCAGATCGGGGAGTTTATAGTACACCAGGGGCTGGCTTACTACTGCTG 839  
 QY 142 GlnValHisPheAspGluGlyLysAlaValTyrrLeuLysLeuAspLeuValAsnGly 161  
 DB 840 CAGGTCCACTTGTATGAGGGAGAGGCTGTCTACCTGAAGCTGACTGTGCTGATGCT 899  
 QY 162 ValLeuAlaLeuArgCysGluGluGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181  
 DB 900 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 959  
 QY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 960 CACCTCCGCTGTCGACAGGAGTCTGGGCTGTGGCCCTGGGCGGAGGAGGCTCTGCGG 1019  
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrrPheGlyLeu 221  
 DB 1020 ATCCGACACCTCCCGGGGCGCCATCAAGGCTGCCCTCTCTCACTCACTTGAGATC 1079  
 QY 222 PheGlnValHis 225  
 DB 1080 TTCACAGTTTCAC 1091

RESULT 8  
 AC069459 203083 bp DNA linear HTG 27-JUN-2001  
 LOCUS AC069459.9  
 DEFINITION Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT  
 SEQUENCE 7 unordered pieces.  
 ACCESSION AC069459 GI:14547768  
 VERSION AC069459  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.

REFERENCE  
 1 (bases 1 to 203083)  
 Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.

TITLE  
 JOURNAL  
 Unpublished

REFERENCE 2 (bases 1 to 203083)  
 AUTHORS Morley, K. C.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-May-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Jun 26, 2001 this sequence version replaced gi:12621364.

COMMENT ----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: MAPO

Center clone name: RP23-168P5

Summary Statistics

Sequencing vector: M13; L08821

Chemistry: Dye-Primer Bodypy: 48% of reads

Chemistry: Dye-terminator Big Dye: 52% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 212648 bases at least Q40

Consensus quality: 212648 bases at least Q30

Estimated insert size: 210656; sum-of-coverage estimation

Quality coverage: 0x in Q20 bases; agarose-gel estimation

Quality coverage: 7.2x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently

consists of 7 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 62152: contig of 62152 bp in length  
 \* 62153 62252: gap of unknown length  
 \* 62253 118772: contig of 56520 bp in length  
 \* 118773 118872: gap of unknown length  
 \* 118873 148924: contig of 30052 bp in length  
 \* 148925 149024: gap of unknown length  
 \* 149025 167231: contig of 18207 bp in length  
 \* 167232 167331: gap of unknown length  
 \* 167332 189907: contig of 22576 bp in length  
 \* 189908 190007: gap of unknown length  
 \* 190008 196537: contig of 6530 bp in length  
 \* 196538 196637: gap of unknown length  
 \* 196638 203083: contig of 6446 bp in length.

FEATURES

SOURCE

1. 203083

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="11"

/clone="RP23-168P5"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others

ORIGIN

Alignment Scores:

Pred. No.: 1.06e-34 Length: 203083  
 Score: 602.50 Matches: 125  
 Percent Similarity: 70.00% Conservative: 1  
 Best Local Similarity: 69.44% Mismatches: 0  
 Query Match: 51.85% Indels: 54  
 DB: 2 Gaps: 1

US-09-245-198a-2 (1-225) x AC069459 (1-203083)

Oy 100 AlagiyValaspGlyThrValserGlyTrpGluGluThrLysIleasnSerSerPro 119  
 Db 43305 TCAGCTGTGATGAGCAGTGTGCTGAGGAAGACCAAAATCAACAGCTCCAGCCCT 43246

Oy 120 LeuArgTyraSPARGlnIleGlyLupheThraValIleArgAlaGlyLeuTyrtyleu 139  
 Db 43245 CTGCGCTACGACCGCCCAATTGGGAATTTACAGTCAATCAGGGCTGACTACTACTG 43186  
 Oy 140 TTTCTT----- 141  
 Db 43185 TACTGTCAAGTAAAGCCTGCTGCTGCTGAGGGCAGAGCAAAAGGTAAAGGAGGAGGA 43126  
 Oy 141 ----- 141  
 Db 43125 CTGCAAGAAATGTTGGGAGAGAGAGACTGTGAGTCAATGAAAGAGCCGTGCTT 43066  
 Oy 142 -----GlnValHisPheAs 146  
 Db 43065 TCAGTACGAGGAGGAGAGAGTGTGATTTGCTTCTCTCTCTCTCTCTCTCTCTCTCT 43006  
 Oy 146 pGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGlyValIleuAlaLeuAr 166  
 Db 43005 TGAGGGAAGAGCTGTCTACTGAACTGAGTCTGCTGAGGAGGAGGAGGAGGAGG 42946  
 Oy 166 gCysLeuGluGluPheSerAlaThrAlaAlaSerProGlyProGlnLeuArgLeuCy 186  
 Db 42945 CTGCTTGAAGAATTTCTAGCCACAGCAAGCTCTCTCTGAGGAGGAGGAGGAGG 42886  
 Oy 186 sGlnValserGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuPr 206  
 Db 42885 CCAGTGTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 42826  
 Oy 206 cTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225  
 Db 42825 CTGGGCTCATCTTAAAGCTGCCCCCTTCTTAACTTGTGACTTTCAGTTCAC 42768

RESULT 9  
 AL603707/c 234182 bp DNA linear ROD 17-NOV-2001  
 LOCUS Mouse DNA sequence from clone RP23-422L16 on chromosome 11,  
 DEFINITION complete sequence.  
 ACCESSION AL603707  
 VERSION AL603707.5 GI:17017790  
 KEYWORDS HNC.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 REFERENCES  
 1. (bases 1 to 234182)  
 Pearce, A.  
 Direct Submission  
 Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
 On Nov 20, 2001 this sequence version replaced gi:16605765.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; Sw.,  
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-422L16 is  
 from the RPI-23 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBAC3.6













```

Db 42506 GGGGCTCTCGCGGCCCCAGGTGACTTGTATGAGGGAGGAGCTGTACTGTAGACT 42447
Qy 155 unapleuLeuValasnglyValleuAlaleuArycysleugluPheseralAthrAl 175
Db 42446 GGACTTGTGTGGTATGAGCGCCCTGGCTGCTCCGTAAGAGTTCTCCGCCACAGC 42387
Qy 175 alaserSerProglyProgluLeuArycysglnValserGlyleuLeuProleuAr 195
Db 42386 GCGCAGCAGCTCGGCGCCAGCTCGCTGCGCAAGTGTGGCTGTGGCTTGGCCCTCGG 42327
Qy 195 gProglySerSerleuArygIleAryThreupProTrAlAlslleuAlalAalAproPh 215
Db 42326 GCGCGGGTCTCTCCCTGGATCGGACCCCTCCCTGGGCCCAATCAGAGCGCCCTT 42267
Qy 215 eLeuthrTyrrhegLyLeuPheglnValHis 225
Db 42266 CCTACTACTTCTGCGACTCTTCAGGTTAAC 42236

RESULT 14
AC119115 138792 bp DNA linear HTG 18-JUL-2002
DEFINITION Rattus norvegicus clone CH230-320N23, *** SEQUENCING IN PROGRESS
AC119115
VERSION *** 32 unordered pieces.
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 138792)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbarta,J., Benton,J., Blinage,K., Blankenburg,K., Bonnih,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinu,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,U., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsoson,E., Kelly,S., Khan,U., King,L., Korvan,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louisedge,H.,
Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okunolu,G.,
Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Taylor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umanal,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Wolley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Direct Submission
Unpublished
2 (bases 1 to 138792)
Worley,K.C.
Direct Submission
Submitted (25-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 138792)
Worley,K.C.
Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20303440.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GUXG
Center clone name: CH230-320N23
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 112720 bases at least Q40
Consensus quality: 116666 bases at least Q30
Consensus quality: 119165 bases at least Q20

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1023: contig of 1022 bp in length
1123: gap of unknown length
1123: contig of 1272 bp in length
2494: gap of unknown length
2495: gap of unknown length
3663: contig of 1169 bp in length
3664: gap of unknown length
3763: gap of unknown length
4879: contig of 1116 bp in length
4880: gap of unknown length
4979: gap of unknown length
6787: contig of 1808 bp in length
6788: gap of unknown length
6887: gap of unknown length
8192: contig of 1305 bp in length
8193: gap of unknown length
8292: gap of unknown length
8293: contig of 1169 bp in length
9461: gap of unknown length
9462: gap of unknown length
9562: contig of 1874 bp in length
11435: contig of 1874 bp in length
11436: gap of unknown length
11436: contig of 2000 bp in length
11536: gap of unknown length
11536: contig of 2000 bp in length
13536: gap of unknown length
13536: contig of 3688 bp in length
17324: gap of unknown length
17324: gap of unknown length
20406: contig of 2963 bp in length
20407: gap of unknown length
20407: gap of unknown length
22445: contig of 2939 bp in length
23445: gap of unknown length
23446: gap of unknown length
23545: gap of 2055 bp in length
25600: gap of unknown length
25701: gap of unknown length
28049: contig of 2349 bp in length
28050: gap of unknown length
28050: gap of unknown length
30474: contig of 2325 bp in length
30475: gap of unknown length
30574: gap of unknown length
30575: contig of 4924 bp in length
35598: gap of unknown length

```

```

* 35599 39284: contig of 3686 bp in length
* 39285 39384: gap of unknown length
* 39385 41970: contig of 2586 bp in length
* 41971 42070: gap of unknown length
* 42071 45659: contig of 3589 bp in length
* 45660 45759: gap of unknown length
* 45760 50332: contig of 4573 bp in length
* 50333 54695: contig of 4263 bp in length
* 54696 58289: contig of 3494 bp in length
* 58290 58390: gap of unknown length
* 58391 63600: contig of 5210 bp in length
* 63601 70305: contig of 6606 bp in length
* 70306 76123: contig of 5718 bp in length
* 76124 76223: gap of unknown length
* 76224 84961: contig of 8738 bp in length
* 84962 93614: contig of 8553 bp in length
* 93615 93714: gap of unknown length
* 93715 103352: contig of 9638 bp in length
* 103353 103452: gap of unknown length
* 103453 110300: contig of 6847 bp in length
* 110301 116180: contig of 5781 bp in length
* 116181 125484: gap of unknown length
* 125485 125584: gap of 9204 bp in length
* 125585 138792: contig of 13208 bp in length.
Location/Qualifiers
1. 138792
/db_xref="taxon:10116"
/clone="CH230-320N23"

```

```

BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 7 23e-30 Length: 138792
Score: 536.50 Matches: 110
Percent Similarity: 67.46% Conservative: 2
Best Local Similarity: 65.09% Mismatches: 4
Query Match: 46.17% Indels: 53
DB: 2 Gaps: 1

```

```

US-09-245-198a-2 (1-225) x AC119115 (1-138792)

```

```

QY 101 GYValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeu 120
|||
Db 137345 GGTCGTGATGGAGACAGTGAAGTGGCTGGAGAACCAAAATCAACACCTCCAGCCCTCTG 137404
QY 121 ArgGlyTrpAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140
|||
Db 137405 CGCATATACCGCCGAGATTGGGGAATTACGGTCATCAGGCGTGGCTACTACTCTGTAC 137464
QY 141 Cys----- 141
|||
Db 137465 TGTCAAGTAAAGCCCTGGGCTCCATGGGTAGACGATGGCTAAGGGAGAAAGCTTGCCA 137524
QY 141 ----- 141
Db 137525 AGAATGCGTGGAGTGGAGAACCTGGGTTTCATGAAGAAGATGCTGGATTTCGATG 137584
QY 142 -----GlnValHisPheAspGlu 147
|||
Db 137585 AGGCGACAGCAGAGCTCGATTGCTGCTCTCTCTCTCCAGGTGCACATTGTGATGAG 137644
QY 148 GLYValAlaValTyrLeuLysLeuAspLeuLeuValAsnGlyValIleuAlaIleuArgCys 167
|||
Db 137645 GCGAAGGCACTTACTTCAAGCTGAGACTGCTGCTGTAATGATGCTGCGCCCTGCGCTGC 137704

```

```

QY 168 LeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArgCysGln 187
|||
Db 137705 CTGGAAGAATTCTCAGCAGACAGCAGCTCTCTCTGGGCCAGCTCCCTGTGTCAG 137764
QY 188 ValSerIleuLeuPheProLeuArgProGlySerSerIleuArgIleuArgThrLeuProTrp 207
|||
Db 137765 GTGCTGGGCTGTGGCTTCGCGACGAGGCTTCCCTTGGATCCCTACCATCACTTGC 137824
QY 208 AlaHisLeuValAlaIleuProPheLeu 216
|||
Db 137825 AGTATATTAAAGCGGCTACTCTGCTT 137851

```

```

RESULT 15
AC126921
LOCUS
DEFINITION
Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
AC126921 GI:21724098
VERSION
AC126921.1
KEYWORDS
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Bos taurus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 153553)
REFERENCE
AUTHORS
Akhtar,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granter,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P.,
Lee-Lin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskerl,B., Masrihan,S.D.,
McCloskey,J.C., McDowell,J., Pegurigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Shanriddop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsugeon,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 153553)
Green,E.D.
Direct Submission
Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Galtersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoemhgrl.nih.gov
----- Project Information
Center clone name: 045D24
Center project name: dd1
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 146066 bases at least Q40
Consensus quality: 147748 bases at least Q30
Consensus quality: 148824 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 152453; sum-of-ctrls
Quality coverage: 8.80x in Q20 bases; agarose-fp
Quality coverage: 8.72x in Q20 bases; sum-of-ctrls
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2312: contig of 2312 bp in length
* 2313 2412: gap of unknown length

```

```

*      2413      5841: contig of 3429 bp in length
*      5842      5941: gap of unknown length
*      5942      8435: contig of 2494 bp in length
*      8436      8535: gap of unknown length
*      8536      15799: contig of 7264 bp in length
*      15800      15899: gap of unknown length
*      15900      25224: contig of 9325 bp in length
*      25225      25324: gap of unknown length
*      25325      32504: contig of 7180 bp in length
*      32505      32604: gap of unknown length
*      32605      40970: contig of 8366 bp in length
*      40971      41070: gap of unknown length
*      41071      56590: contig of 15520 bp in length
*      56591      56690: gap of unknown length
*      56691      73770: contig of 17079 bp in length
*      73770      73869: gap of unknown length
*      73870      90859: contig of 16990 bp in length
*      90860      90959: gap of unknown length
*      90960      111428: contig of 20469 bp in length
*      111429      111528: gap of unknown length
*      111529      153553: contig of 42025 bp in length.
FEATURES
Source
1..153553
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-45D24"
/clone_1lb="RP42"
1..2312
/note="assembly-fragment"
2413..5841
/note="assembly-fragment"
clone_end:r7
vector_side:left"
5942..8435
/note="assembly-fragment"
8536..15799
/note="assembly-fragment"
15900..25224
/note="assembly-fragment"
25325..32504
/note="assembly-fragment"
32605..40970
/note="assembly-fragment"
41071..56590
/note="assembly-fragment"
56691..73770
/note="assembly-fragment"
73870..90859
/note="assembly-fragment"
90960..111428
/note="assembly-fragment"
111529..153553
/note="assembly-fragment"
clone_end:SP6
vector_side:right"

```

```

BASE COUNT   39178 a 37393 c 36616 g 39259 t 1107 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      2,81e-27      Length:      153553
Score:          503.00      Matches:      111
Percent Similarity: 58.59%      Conservative: 5
Best Local Similarity: 56.06%      Mismatches: 17
Query Match:     43.29%      Indels:      66
DB:              2          Gaps:      1

```

```

US-09-245-198a-2 (1-225) x AC126921 (1-153553)

```

```

QY      93      ProGlyGlnAspGlyAlaGlnIaGlyValaSpGlyThrValSerGlyTrpGluGluThr 112
DB      26848  CCTGCTCTGAAACACCCACATTCAGGTGTGACGCGGTGAGTGTGCGGAGAGGCC 26907
QY      113  LysIleAsnSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIle 132

```

```

DB      26908  AAATATACAGCTCCCAACCCCTGGCTATGACTGCGACCGGGCAATTTCAGTCACCC 26967
QY      133      ArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal----- 143
DB      26968  CGGGCTGGGCTGACTACTGACTGACTGAGT- AAGCCCACTGCGCTCCAGGGTAAAG 27026
QY      143      ----- 143
DB      27027  CCGGAACGTAGAGAGAAAGGCTGGCTTCGGGGTGGGGCAAGTTAAAGTGGGAGGG 27086
QY      143      ----- 143
DB      27087  GAGCGTGGGTTTGGGCTGAGAGAGAGCTTGGGCTCTAAGAGACACTGAGATGAAGCCC 27146
QY      144      -----HisPheAspGlu 147
DB      27147  AGGGCCAGCAGAGAGGCTGGACTCCGCCCTCCCTGCCCTCCAGCTGCACCTTGTGATGAG 27206
QY      148  GlyLysAlaValTyrLeuTyrLysIleuAspLeuValAsnGlyValLeuAlaLeuArgCys 167
DB      27207  GGGAGGCTGCTACTGAACTGAGCTGCTGTGTGATGACACGCTGGCCCTGCGCTGC 27266
QY      168  LeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeuCysGln 187
DB      27267  CTGAGAGAAATTCCTCGGCACACTGGGCGGCGAGTCCCTGGGCGCCAGCTCCGCTCTGCCAA 27326
QY      188  ValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTrp 207
DB      27327  GTGTCAAGGCTCTTGTGGCTGGCGGCGAGGCTCTTCCCTGGGATCCGACCCCTCCAG 27386
QY      208  AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225
DB      27387  ACCCACTCAAGGCTCCCTTCCTCCTCACCCTACTTTGGACTCTTCCAAAGTTTCAAC 27440

```

```

Search completed: March 31, 2003, 04:35:32
Job time : 3276.48 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:45 ; Search time 42.4361 Seconds  
(without alignments)  
706.506 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSTGLALACGLLVVSL.....PMAHLKAPFLTYRGLFQVH 225

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162	100.0	225	19	AAW47524	Mus musculus tumou
2	1162	100.0	225	21	AAW47527	Amino acid sequenc
3	1089	93.7	211	20	AAW3591	Mouse TNF3 protei
4	1020	87.8	249	19	AAW29745	TNF related endoth
5	1020	87.8	249	20	AAW09369	Human tumour necro
6	1020	87.8	249	21	AAW07526	Amino acid sequenc
7	1020	87.8	249	21	AAW5338	Human PRO207 antit
8	1020	87.8	249	22	AAW00891	Human TRPA (TNF r
9	1020	87.8	249	23	AAW06129	Human PRO207 polyp
10	1020	87.8	284	19	AAW47525	Homo sapiens tumou

11	951	81.8	208	20	AAW3590
12	951	81.8	273	22	AAU03499
13	742	63.9	189	19	AAW29746
14	742	63.9	189	22	AAW00892
15	721	62.0	146	22	AAE00895
16	111	9.6	406	23	AAU77717
17	110.5	9.5	409	23	AAU77718
18	107.5	9.3	325	22	ABW67553
19	106.5	9.2	294	19	AAW69356
20	106.5	9.2	294	19	AAW68292
21	106.5	9.2	294	22	AAE08737
22	106.5	9.2	294	22	AAE04425
23	106.5	9.2	294	22	AAE01992
24	103	8.9	220	22	AAW62340
25	102.5	8.8	316	19	AAW83194
26	102.5	8.8	316	19	AAW83017
27	102.5	8.8	316	19	AAW59654
28	102.5	8.8	316	20	AAV17874
29	102.5	8.8	316	21	AAV91024
30	102.5	8.8	316	21	AAW84418
31	102.5	8.8	316	21	AAW84419
32	102.5	8.8	316	23	AAU78289
33	99	8.5	234	22	AAW62339
34	97	8.3	255	22	AAW86311
35	96.5	8.3	409	23	AAU77716
36	92	7.9	261	23	AAW49225
37	91	7.8	182	18	AAW08128
38	91	7.8	294	18	AAW09123
39	90.5	7.8	211	21	AAV58216
40	90.5	7.8	260	21	AAV58215
41	90.5	7.8	318	22	AAW82092
42	90.5	7.8	531	10	AAW91776
43	90	7.7	182	18	AAW24011
44	90	7.7	182	18	AAW24013
45	90	7.7	261	18	AAW09115

#### ALIGNMENTS

RESULT 1	
AAW47524	AAW47524 standard; Protein: 225 AA.
ID	AAW47524 standard; Protein: 225 AA.
AC	AAW47524;
DT	21-JUL-1998 (first entry)
XX	
DE	Mus musculus tumour necrosis factor related ligand (TRELL).
XX	
XX	TRELL: tumour necrosis factor related ligand; tnfr: treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection.
XX	
OS	Mus musculus.
XX	
EH	Key
FT	Domain
XX	
PN	Location/Qualifiers
XX	
PD	12-FEB-1998.
XX	
PF	07-AUG-1997; 97WO-US13945.
XX	
PR	18-MAR-1997; 97US-0040820.
PR	07-AUG-1996; 96US-0023541.
PR	18-OCT-1996; 96US-0028515.
XX	
PA	(BIOT ) BIOGEN INC.
PA	(UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX	

PI Browning JL, Chicheportliche Y;  
 XX  
 DR WPI: 1998-145619/13.  
 DR N-PSDB; AAV18599.  
 XX  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 12; Pages 48-50; 69pp; English.  
 XX  
 CC The sequence is that of mouse tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, auto-immune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. It's coding sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC auto-immune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of a gene encoding TRELL  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 CC  
 XX  
 SQ Sequence 225 AA:  
 Query Match 100.0%; Score 1162; DB 19; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VLSIGLALACGLLLVYVSGSWATLSAQPSELTAEEDRREPPELNPQTESQDVVPF 60  
 DB 1 VLSIGLALACGLLLVYVSGSWATLSAQPSELTAEEDRREPPELNPQTESQDVVPF 60  
 QY 61 LEQVLRPRRSAPKGRKARPRRAIAHVEVHPRPGQDAQAGVDGTSGWEETKINSSPL 120  
 DB 61 LEQVLRPRRSAPKGRKARPRRAIAHVEVHPRPGQDAQAGVDGTSGWEETKINSSPL 120  
 QY 121 RYDQIQIEFTYIRAGLYLYLCQVHFDEGKAVYLLKLDLLVNGVIALRCLEEFSAATASSPG 180  
 DB 121 RYDQIQIEFTYIRAGLYLYLCQVHFDEGKAVYLLKLDLLVNGVIALRCLEEFSAATASSPG 180  
 QY 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225  
 DB 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225  
 RESULT 2  
 AAB07527 standard; protein: 225 AA.  
 XX  
 AC AAB07527;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soluble recombinant murine TWEAK protein.  
 XX  
 KW TWEAK protein; immunological disorder; immune response; inflammation;  
 KW TWEAK blocking agent; auto-immune disease; organ transplant rejection;  
 KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200042073-A1.  
 XX  
 PD 20-JUL-2000.  
 XX

PF 14-JAN-2000; 2000WO-US01044.  
 XX  
 XX 15-JAN-1999; 99US-0116168.  
 XX  
 PA (BIO) BIOGEN INC.  
 XX  
 PI Renmert P;  
 XX  
 DR WPI: 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease.  
 XX  
 PS Disclosure; Fig 1; 45pp; English.  
 XX  
 CC The present sequence represents a TWEAK protein. The specification  
 CC describes a method for preventing or treating an immunological  
 CC disorder and/or inhibiting an immune response in an animal. The  
 CC method comprises administering a TWEAK blocking agent. The method may  
 CC be used for preventing and treating immune disorders associated with  
 CC inappropriate expression and/or activity of TWEAK. These disorders  
 CC include autoimmune diseases, acute and chronic inflammation, organ  
 CC transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell  
 CC malignancies, septic and other forms of shock, loss of immune  
 CC responsiveness (as seen in human immunodeficiency virus (HIV)  
 CC infections) and failure of the immune response to tumour growth.  
 CC  
 XX  
 SQ Sequence 225 AA:  
 Query Match 100.0%; Score 1162; DB 21; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;  
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VLSIGLALACGLLLVYVSGSWATLSAQPSELTAEEDRREPPELNPQTESQDVVPF 60  
 DB 1 VLSIGLALACGLLLVYVSGSWATLSAQPSELTAEEDRREPPELNPQTESQDVVPF 60  
 QY 61 LEQVLRPRRSAPKGRKARPRRAIAHVEVHPRPGQDAQAGVDGTSGWEETKINSSPL 120  
 DB 61 LEQVLRPRRSAPKGRKARPRRAIAHVEVHPRPGQDAQAGVDGTSGWEETKINSSPL 120  
 QY 121 RYDQIQIEFTYIRAGLYLYLCQVHFDEGKAVYLLKLDLLVNGVIALRCLEEFSAATASSPG 180  
 DB 121 RYDQIQIEFTYIRAGLYLYLCQVHFDEGKAVYLLKLDLLVNGVIALRCLEEFSAATASSPG 180  
 QY 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225  
 DB 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225  
 RESULT 3  
 AAW93591  
 ID AAW93591 standard; protein: 211 AA.  
 XX  
 AC AAW93591;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Mouse TNRL3 protein.  
 XX  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX



```
XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;  
KM NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
KM cancer.  
XX  
OS Homo sapiens.  
XX  
PN MO9919490-A1.  
XX  
PD 22-APR-1999.  
XX  
PF 09-OCT-1998; 98WO-US21407.  
XX  
PR 17-DEC-1997; 97US-0069862.  
PR 10-OCT-1997; 97US-0062037.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Marsters SA, Pictl R;  
XX  
DR WPI; 1999-287982/24.  
DR N-PSDB; AAX56000.  
XX  
PT New human Apo-3 ligand (a tumour necrosis factor) homologue  
XX  
PS Claim 1; Fig 1; 74pp; English.  
XX  
CC The present sequence represents a human tumour necrosis factor (TNF)  
CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in  
CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and  
CC to induce JNK/SAPK-dependent responses in mammalian cells.  
XX  
SQ Sequence 249 AA;  
XX  
Query Match 87.8%; Score 1020; DB 20; Length 249;  
Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
XX  
QY 2 LSLGLALACGLLLVYVSLGSMATLSAQPSELTAEDERRPEELNPQTEESQDVVFL 61  
DB 26 LGGLALACGLLLVYVSLGSRASLSAQEPQAEELVAEEDDPSELNPQTEESQDPAFL 85  
QY 62 EQLVPRRSAPKGRKARRRAIAAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSPLR 121  
DB 86 NRLVPRRSAPKGRKARRRAIAAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSPLR 145  
QY 122 YDRQIGFTYTRAGLYLYCOVHDEGKAVYLYKLDLVNGLVLAIRCLEEFSAATASSGPR 181  
DB 146 YNRQIGFTYTRAGLYLYCOVHDEGKAVYLYKLDLVNGLVLAIRCLEEFSAATASSGPR 205  
QY 182 QLRLQGVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYFGLFQVH 225  
DB 206 QLRLQGVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYFGLFQVH 249  
XX  
RESULT 6  
AAB07526  
ID AAB07526 standard; protein; 249 AA.  
XX  
AC AAB07526;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Amino acid sequence of a soluble recombinant human TWEAK protein.  
XX  
KW TWEAK protein; immunological disorder; immune response; inflammation;  
KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
KW Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200042073-A1.
```

```
XX 20-JUL-2000.  
PD 14-JAN-2000; 2000WO-US01044.  
XX  
PF 15-JAN-1999; 99US-0116168.  
XX  
PR (BIOJ ) BIOGEN INC.  
XX  
PA Renner P;  
XX  
PI WPI; 2000-476036/41.  
XX  
DR Preventing and treating immune responses using modulators, especially  
XX antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
XX treating e.g. inflammation and graft versus host disease -  
XX Disclosure; Fig 1; 45pp; English.  
XX  
CC The present sequence represents a TWEAK protein. The specification  
CC describes a method for preventing or treating an immunological  
CC disorder and/or inhibiting an immune response in an animal. The  
CC method comprises administering a TWEAK blocking agent. The method may  
CC be used for preventing and treating immune disorders associated with  
CC inappropriate expression and/or activity of TWEAK. These disorders  
CC include autoimmune diseases, acute and chronic inflammation, organ  
CC transplant rejection, graft-versus-host disease (GVHD), lymphoid cell  
CC malignancies, septic and other forms of shock, loss of immune  
CC responsiveness (as seen in human immunodeficiency virus (HIV)  
CC infections) and failure of the immune response to tumour growth.  
XX  
SQ Sequence 249 AA;  
XX  
Query Match 87.8%; Score 1020; DB 21; Length 249;  
Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
XX  
QY 2 LSLGLALACGLLLVYVSLGSMATLSAQPSELTAEDERRPEELNPQTEESQDVVFL 61  
DB 26 LGGLALACGLLLVYVSLGSRASLSAQEPQAEELVAEEDDPSELNPQTEESQDPAFL 85  
QY 62 EQLVPRRSAPKGRKARRRAIAAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSPLR 121  
DB 86 NRLVPRRSAPKGRKARRRAIAAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSPLR 145  
QY 122 YDRQIGFTYTRAGLYLYCOVHDEGKAVYLYKLDLVNGLVLAIRCLEEFSAATASSGPR 181  
DB 146 YNRQIGFTYTRAGLYLYCOVHDEGKAVYLYKLDLVNGLVLAIRCLEEFSAATASSGPR 205  
QY 182 QLRLQGVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYFGLFQVH 225  
DB 206 QLRLQGVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYFGLFQVH 249  
XX  
RESULT 7  
AAV95338  
ID AAV95338 standard; protein; 249 AA.  
XX  
AC AAV95338;  
XX  
DT 25-SEP-2000 (first entry)  
XX  
DE Human PRO207 antitumour protein.  
XX  
KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
KW central nervous system cancer; melanoma; leukaemia; neoplasm.  
XX  
OS Homo sapiens.  
XX  
PN Key Location/Qualifiers  
FT Peptide 1..40
```

Seq	Sequence	249 AA;	87.8%;	Score 1020;	DB 21;	Length 249;
Query	Match		88.8%;	Pred. No. 9.5e-98;		
Best	Local Similarity					
Matches	199;	Conservative 9;	Mismatches 16;	Indels 0;	Gaps 0;	
QY	2	LSLGALACDLILLVVSLSGSMATISAOPEQDELTAADREPPPLNQTDESDVYFLL	61			
DB	26	LGLGLALACDLILLVVSLSGRASISAOPEQDELTAADREPPPLNQTDESDVYFLL	85			
QY	62	EOLVPRRSAPKGRKAPRRRAIAAHYEVHPPRGDGAAGVDGTVSGWEETKINSSPLR	121			
DB	86	NRLVPRRSAPKGRKTRARRRAIAAHYEVHPPRGDGAAGVDGTVSGWEETKINSSPLR	145			
QY	122	YDRQIGETFTVIRAGLYLYLCVHFEDEGRAVYLIKLDLVNGVALRCLEEFSAATASSGP	181			
DB	146	YNRQIGETFTVIRAGLYLYLCVHFEDEGRAVYLIKLDLVNGVALRCLEEFSAATASSGP	205			
QY	182	QLRLCVSGLLPLRPGSSLRIRTPMALLKAAPFLITYGLQVH	225			
DB	206	QLRLCVSGLLPLRPGSSLRIRTPMALLKAAPFLITYGLQVH	249			
RESULT 8						
ID	AAE00891					
AC	AAE00891	standard; Protein: 249 AA.				
XX	AAE00891;					
DT	04-JUL-2001	(first entry)				
XX						
DE	Human TREPA (TNF related endothelium proliferative agent).					
XX						
KW	Human; tumour necrosis factor; TNF; angiogenesis; wound healing;					
KW	TREPA; TNF related endothelium proliferative agent; tumour; metastasis;					
KW	grafting; vulnetary.					
XX						
OS	Homo sapiens.					
XX						
FT	Key	Location/Qualifiers				
FT	Domain	98..249				
FT		/label= Extracellular_domain				
XX						
PN	US6207642-B1.					
XX						
PD	27-MAR-2001.					
XX						
PF	26-JUN-1998;	98US-0105343.				
XX						
PR	12-FEB-1997;	97US-0798692.				
PR	10-FEB-1998;	98US-0021706.				
XX						
PA	(ABBO ) ABBOTT LAB.					
XX						
PI	Wiley SR;					
XX						
DR	WPI: 2001-280760/29.					
DR	N-PSDB: AAD04350.					
XX						
XX						
CC	Claim 1; Column 75-76; 53pp; English.					
CC	The present invention relates to extracellular signal molecules,					
CC	particularly members of tumour necrosis factor (TNF) family molecules					
CC	designated as TREPA (TNF related endothelium proliferative agent).					
CC	Soluble biologically active TREPA are used to treat TREPA-associated					
CC	diseases, tumours or metastases. TREPA is used for inducing angiogenesis					
CC	in human for promoting wound healing and for vascularising grafted					
CC	tissue for successful grafting and to promote tissue grafts.					



DR WP1: 1998-145619/13.  
 DR N-PSDB: AAV18600.  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 12: Pages 50-51; 69pp; English.  
 CC The sequence is that of human tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. It's coding sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of prepare probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 XX  
 SO Sequence 284 AA:  
 Query Match 87.8%; Score 1020; DB 19; Length 284;  
 Best Local Similarity 88.8%; Pred. No. 1.1e-97;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 LSLGLALACGLLVVSLGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPL 61  
 Db 61 LGLGLALACGLLVVSLGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPL 120  
 QY 62 EOLVPRRSAPKGRKARRRAIAHYEVHPRPGDGAQVDTVSGMEAKINSSPLR 121  
 Db 121 NRVPRRSAPKGRKARRRAIAHYEVHPRPGDGAQVDTVSGMEAKINSSPLR 180  
 QY 122 YDRQIGFEFTVIRAGLYLYLCQVHFEDEKAVYLLKIDLNVGVLAIRCLSEFSATASPGP 181  
 Db 181 YNRQIGFEFTVIRAGLYLYLCQVHFEDEKAVYLLKIDLNVGVLAIRCLSEFSATASPGP 240  
 QY 182 QLRLCQVSGLLPLRPGSSLRIRTPMAHLKAAPFLTYFGLFQVH 225  
 Db 241 QLRLCQVSGLLPLRPGSSLRIRTPMAHLKAAPFLTYFGLFQVH 284  
 RESULT 11  
 ID AAW93590 standard; Protein: 208 AA.  
 XX AAW93590:  
 XX 18-JUN-1999 (first entry)  
 DE Human TNRL3 protein.  
 XX  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; ApO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW ApO6; ApO8; ApO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX MO9911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX

PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIV ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WP1: 1999-205191/17.  
 DR N-PSDB: AAX23424.  
 XX  
 PT New Tumour Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Claim 40; Fig 13A; 156pp; English.  
 CC This invention describes isolated Tumour Necrosis Factor (TNF) family  
 CC receptor polypeptides: ApO4, ApO6, ApO8 and ApO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. ApO4 is useful for diagnosing prostate cancer  
 CC by determining levels of ApO4 in an individual. Prostate cancer can also  
 CC be treated using ApO4 selective binding agents linked to a therapeutic  
 CC moiety. ApO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. ApO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in ApO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using ApO4 polypeptides/active  
 CC fragments and ApO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. ApO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. ApO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 SO Sequence 208 AA:  
 Query Match 81.8%; Score 951; DB 20; Length 208;  
 Best Local Similarity 88.9%; Pred. No. 1.1e-90;  
 Matches 185; Conservative 8; Mismatches 15; Indels 0; Gaps 0;  
 QY 18 VSLGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPLFQVLRPRRSAPKGRKA 77  
 Db 1 VSLGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPLFQVLRPRRSAPKGRKT 60  
 QY 78 RPRRAIAHYEVHPRPGDGAQVDTVSGMEAKINSSPLRDRQIGFEFTVIRAGLY 137  
 Db 61 RARRRAIAHYEVHPRPGDGAQVDTVSGMEAKINSSPLRKYRQIGFEFTVIRAGLY 120  
 QY 138 YLYCQVHFEDEKAVYLLKIDLNVGVLAIRCLSEFSATASPGQLRLCQVSGLLALRPG 197  
 Db 121 YLYCQVHFEDEKAVYLLKIDLNVGVLAIRCLSEFSATASPGQLRLCQVSGLLALRPG 180  
 QY 198 SSLRIRTPMAHLKAAPFLTYFGLFQVH 225  
 Db 181 SSLRIRTPMAHLKAAPFLTYFGLFQVH 208  
 RESULT 12  
 ID AAU03499 standard; Protein: 273 AA.  
 XX AAU03499:  
 XX 26-SEP-2001 (first entry)  
 DE TWEAK extracellular domain-containing fusion protein.  
 XX  
 DE TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 XX

KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KM rubrosis; uveitis; macular degeneration; arthritis; rheumatism;  
 KM corneal graft neovascularisation; psoriasis; metastatic condition;  
 KM malignant tumour; sarcoma; carcinoma; benign tumour; hemophilic joint;  
 KM preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KM peripheral atherosclerosis; PDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO200145730-A2.  
 XX  
 XX 28-JUN-2001.  
 PD  
 PF 19-DEC-2000; 2000WO-US34755.  
 XX  
 XX 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 XX (IMMUNEX ) IMMUNEX CORP.  
 PA Willey SR.  
 PI  
 XX  
 DR N-PSDB; AAS03964.  
 XX  
 PS Example 1; Page 41; 46pp; English.  
 CC The sequence represents a fusion protein encoded by the  
 CC expression vector PDC409-L2-TWEAK. The fusion protein comprises a  
 CC growth hormone leader, a leucine zipper multimerisation domain, and  
 CC the extracellular domain of human TWEAK. The fusion protein was  
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 CC clones from a COS cell human cDNA library. The TWEAK protein is  
 CC a member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and  
 CC develop TWEAKR agonists and antagonists for the modulation of  
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
 CC The disorders mediated by angiogenesis include ocular disorders  
 CC characterised by ocular neovascularisation such as diabetic retinopathy,  
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 CC retrolental fibroplasia, rubrosis, uveitis, macular degeneration and  
 CC corneal graft neovascularisation, and inflammatory diseases such as  
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
 CC malignant and metastatic conditions such as sarcomas and carcinomas,  
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 CC  
 XX  
 SO Sequence 273 AA:  
 Query Match 81.8%; Score 951; DB 22; Length 273;  
 Best Local Similarity 88.9%; Pred. No. 1.6e-90;  
 Matches 184; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 139 LYCOVHDESKAVYELKLDLVNGVLAARCLSEFSATASSPGQRLRCQVSGLLPRPGS 198  
 DB LYCVHDESKAVYELKLDLVNGVLAARCLSEFSATASSPGQRLRCQVSGLLPRPGS 246  
 QY 199 SLRIRLPMANHLKAAPFLYFGLFQVH 225  
 DB 247 SLRIRLPMANHLKAAPFLYFGLFQVH 273  
 RESULT 13  
 AAM29746  
 ID AAM29746 standard; Protein: 189 AA.  
 XX  
 AC AAM29746;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE TNF related endothelium proliferative agent protein 2.  
 XX  
 KM TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
 KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9835061-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 12-FEB-1998; 98WO-US02859.  
 XX  
 PR 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Willey SR.  
 XX  
 DR WPI; 1998-447255/38.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 XX  
 PS Claim 16; Page 125-6; 142pp; English.  
 XX  
 XX The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 XX  
 SO Sequence 189 AA:  
 Query Match 63.9%; Score 742; DB 19; Length 189;  
 Best Local Similarity 92.8%; Pred. No. 5.7e-69;  
 Matches 142; Conservative 3; Mismatches 8; Indels 0; Gaps 0;





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:59:32 ; Search time 2136.97 seconds  
(without alignments)  
8651.954 Million cell updates/sec

Title: US-09-245-198a-1

Perfect score: 1168

Sequence: 1 ggtgctgagcctggccttg.....ataatcatgattctcttc 1168

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues 32308132

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estcpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	883.8	75.7	1033	11	AK020909 Mus muscu
2	625.8	53.6	918	12	BF577781 603092080
3	519.4	44.5	731	13	BI871711 603395825
4	507.4	43.4	728	13	BI870393 603395641
5	504	43.2	561	10	AM763237 ur70d09.y
6	497	42.6	533	10	BE628951 uu30c03.y

7	489.6	41.9	650	12	BG404836
8	488.8	41.8	687	14	BO208433
9	481.2	41.2	584	10	AM917574
10	480.4	41.1	777	13	BI819200
11	473	40.1	834	13	BI766766
12	467.8	40.5	948	14	BO707185
13	456.4	39.1	828	13	BI596681
14	437.8	37.5	471	9	AA221610
15	421.6	36.1	963	14	BO671259
16	419.2	35.9	940	14	BO884231
17	406.4	34.8	1071	14	BM921213
18	390.6	33.4	445	9	AA870722
19	366.6	31.4	951	14	BO674188
20	360	30.8	360	10	BE654876
21	353	30.2	405	9	AT854476
22	331	28.3	367	13	BI111534
23	325.4	27.9	376	9	AA792068
24	309	26.5	474	13	BI965174
25	306	26.2	315	12	BF466521
26	299.4	25.6	538	12	BF821434
27	297.2	25.4	400	9	AT152313
28	292.8	25.1	493	10	BE307031
29	289.6	24.8	433	12	BF283688
30	289.4	24.8	542	12	BF041509
31	285.2	24.4	443	12	BG378802
32	284.4	24.3	785	13	BI762908
33	283.8	24.3	404	12	BF044430
34	278.4	23.8	897	13	BI730298
35	274.4	23.5	412	9	AT101416
36	271	23.2	581	13	BI738634
37	270.4	23.2	402	12	BF410871
38	270.2	23.1	292	10	BB268794
39	258.2	22.1	558	13	BM484863
40	255.2	21.8	394	12	BG376757
41	246.4	21.1	298	14	BM688946
42	243.6	20.9	371	13	BI300370
43	242.4	20.8	441	13	BI967060
44	242.2	20.7	374	9	AA800970
45	240	20.5	240	10	AM764050

## ALIGNMENTS

RESULT 1	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
LOCUS	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
DEFINITION	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
ACCESSION	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
VERSION	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
KEYWORDS	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
SOURCE	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
ORGANISM	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
REFERENCE	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
AUTHORS	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
TITLE	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
JOURNAL	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
MEDLINE	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
PUBMED	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
REFERENCE	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
AUTHORS	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
TITLE	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
JOURNAL	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
MEDLINE	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002
PUBMED	AK020909	1033 bp	muscle	linear	HTC 19-JAN-2002



Db	422	CAGTTTCATGAGGGGCCCTGTCTCGCCAGATTCCTTAACTTTCCCTGGCTCCAGGAC	481
Oy	728	ATCACCACACCTCCCTACCCCACCCCCCACTCTCCACCCCCCTCGCTCTTGCTCCAG	787
Db	482	ATCACCAACACTTTCCTACCCCAACCCCCCACTCTCCACCCCCCTGAAGCTGCTTGCTGCAG	541
Oy	788	TCTCTCT-CTCTCTCTCAAAGGAGCAGCAGAGCTTGTCACATG-TTTCACAATCCACAGAGCT	845
Db	542	TCTCTCTCTCTCTCTCAAAGGAGCAGAGACTTGTCACATGTTTTCCATTCCACAGAGCT	601
Oy	846	ATCTTGTCTCTTC-TTAACATTCCTCATCCGACACACAATATCCACTCTACCTAGCTCCAA	904
Db	602	ATCTTGTCTCTCTTTAACAATCCATCCACACACAATATCCACTCTACCTCCAAA	661
Oy	905	GCCCCCTACTTATCCCTGACTCCGCCACCCACTACCCGACGCACTGTTATATGACTTTGT	964
Db	662	GCCCCCTACTTATCCCTGACTCCGCCACCCACTACCCGACGCACTGTTATATGACTTTGT	721
Oy	965	GCACCAAGCACTGAGATGGGCTGGACCTGTGGGACAGAAGCCAGAGAACCTTGGGACTAGG	1024
Db	722	GCACCAAGCACTGAGATGGGCTGGACCTGTGGGACAGAAGCCAGAGAACCTTGGGACTAGG	781
Oy	1025	CCAGAAAGTTCACATGTGAGGGGGGAAGAGCTGGGGGACAAAGTCTCTCCCTGGATCCCGT	1084
Db	782	CCAGAAAGTTCACACTGTGAGGGGGGAAGAGCTGGGGGACAAAGTCTCTCCCTGGATCCCGT	841
Oy	1085	GGATTTTGAAGAAAGATACTATTTTATTTATTTATTTGTGACAAAAATGTTAAATGATATTAA	1144
Db	842	GGATTTTGAAGAAAGATACTATTTTATTTATTTATTTGTGACAAAAATGTTAAATGATATTAA	901
Oy	1145	GAGATAAATCATGATTTCTCTTC	1168
Db	902	GAGATAAATCATGATTTCTCTTC	925

```

RESULT 2
LOCUS      BF577781
DEFINITION BF577781      918 bp      mRNA      linear      EST 12-DEC-2000
602092080F01 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4206595 5'
ACCESSION  BF577781
VERSION     BF577781.1  GI:11651493
KEYWORDS
SOURCE
ORGANISM   house mouse.
            Mus musculus
            Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria (to 918)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapds@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LMNL at:
            http://image.llnl.gov
            Plate: LLM9767 row: p column: 20
            High quality sequence start: 17
            High quality sequence stop: 724.
            Location/Qualifiers
            1..918
FEATURES
source

```

FEATURES	SOURCE
Location/Qualifiers	1. .918
/organism="Mus musculus"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone_image="IMAGE:4206595"	
/clone_lib="NCI CGAP Co24"	
/lab_host="DH10B (T1 phage-resistant)"	
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT	

Average insert size 1.6 kb. Constructed by life Technologies. Note: this is a NCI-CGAP library.	
BASE COUNT	153 a 292 c 268 g 204 t 1 others
ORIGIN	
Query Match	53.6% Score 625.8; DB 12; Length 918;
Best Local Similarity	99.2%; Pred. No. 7.2e-152;
Matches 660; Conservative	0; Mismatches 2; Indels 3; Gaps 3

OY	22	GCTGGCCTGACCCTTGGACCTTCTGCTGGTCGTGTACAGCCTGGGGAGACGTGGGCAACGCTGTC	81
Db	1	GCTGGCCTGACCCTTGGACCTTCTGCTGGTCGTGTACAGCCTGGGGAGACGTGGGCAACGCTGTC	60
OY	82	TGCCAGGAGACCTTCTCAGGAGAGCTGCACAGACAGACCGCCGGAGCCCTCTGAACT	141
Db	61	TGCCAGGAGACCTTCTCAGGAGAGCTGCACAGACAGACCGCCGGAGCCCTCTGAACT	120
OY	142	GAATCCCCCAGACAGAGGAAAGCCAGAGATGTGGACTTTCTTGGAAACAACCTAGTCCGGCC	201
Db	121	GAATCCCCCAGACAGAGGAAAGCCAGAGATGTGGACTTTCTTGGAAACAACCTAGTCCGGCC	180
OY	202	TGGAAGAAGTGCCTTAAAGGCGGAGAGCGCGCCCTCGCCGAGCTATTGCAAGCCCATTA	261
Db	181	TGGAAGAAGTGCCTTAAAGGCGGAGAGCGCGCGG -CTGCCGAGCTATTGCAAGCCCATTA	239
OY	262	TGAGGTTACCTCTCGGCCAGAGACAGATGGACACACAAGCAGAGTGTGGATGGGACAGTAGAG	321
Db	240	TGAGGTTACCTCTCGGCCAGAGACAGATGGACACACAAGCAGAGTGTGGATGGGACAGTAGAG	299
OY	322	TGGCTGGGGAAGAGACCAAAATCAACAGAGCTCCAGACCCCTCTGCGCTACAGACCGGCAGATTAG	381
Db	300	TGGCTGGGGAAGAGACCAAAATCAACAGAGCTCCAGACCCCTCTGCGCTACAGACCGGCAGATTAG	359
OY	382	GGAATTTACAGTCATCAGAGGCTTGGGCTCTACTACTCTGTACTGTACAGGTGCACTTTGATGA	441
Db	360	GGAATTTACAGTCATCAGAGGCTTGGGCTCTACTACTCTGTACTGTACAGGTGCACTTTGATGA	419
OY	442	GGGAAGAGCTGCTACCTCAACCTGAGACTTGGTGGGAACGGGTGTCTGGCCCTGGGCTG	501
Db	420	GGGAAGAGCTGCTACCTCAACCTGAGACTTGGTGGGAACGGGTGTCTGGCCCTGGGCTG	479
OY	502	CCTGGAAGAAATTCAGCCACAGACAGCAAGCTCTCTCTGGGGCCCAAGCTCCGTTTGTGCCA	561
Db	480	CCTGGAAGAAATTCAGCCACAGACAGCAAGCTCTCTCTGGGGCCCAAGCTCCGTTTGTGCCA	539
OY	562	GGTGTCTGGGCGTGTGGCGCTGCGGCGAGAGGCTTCCCTTGGGATCGGACCGACCTCCCGTG	621
Db	540	GGTGTCTGGGCGTGTGGCGCTGCGGCGCA -GGTCTTCCCTTGGGATCGGACCGACCTCCCGTG	598
OY	622	GGCATCATCTTAAAGCTGGCCCTTCTCTAACACTACTTGTGGACTCTTTCAAAGTTCACTGAGG	681
Db	599	GGCATCATCTTAAAGCTG -CCCTTCTCTAACACTTGTGGACTCTTTCAAAGTTCACTGAGG	657
OY	682	GGCCT 686	
Db	658	GGCCT 662	

RESULT 3				
B1871711	B1871711	731 bp	mRNA	linear EST 11-OCT-2001
LOCUS	603395825F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405478 5'			
DEFINITION	mRNA sequence.			
ACCESSION	B1871711			
VERSION	B1871711.1 GI:16045386			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 731)			
	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
	National Institutes of Health, Mammalian Gene Collection (MGC)			



Db	241	AGCCCATTTATGAAGTTCATCCACAGCACTGGACAGGACGAGCGAGGACAGGTGTGGACGG	3000
Qy	313	GACAGTAGTAGTGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTTCGCGTACGACCG	3727
Db	301	GACAGTAGTAGTGCTGGGAGGAGGACAGAAATCAACAGCTCCAGCCCTTCGCGCTACAACCG	3606
Qy	373	CCAGATTTGGGGAAATTTCACAGTCATCAGGGCTGGGCTCTACTACCTGTACTGTCAAGGTGCA	4322
Db	361	CCAGATCGGGAGATTTCATAGCACACCCGGGCTGGGCTCTACTACCTGTACTGTCAAGGTGCA	4200
Qy	433	CTTTGATGAGGAGAAAGCGCTGTCTACCTGAAAGCTGGACTCTCTGGTGAACGGTGTCTGGC	4922
Db	421	CTTTGATGAGGAGGAAAGCGCTGTCTACCTGAAAGCTGGACTCTCTGGTGAATGGTGTCTGGC	4800
Qy	493	CTTTCGGCTGCCGTGGAAATAATTCCTAGCCACAGCAGCAAGCTCTTCCTTGGGCCCCAGCTCG	5522
Db	481	CTTTCGGCTGCCGTGGAGGAATTCCTAGCCACAGCTGGGCGCAGTTCCCTCGGGCCCCAGCTCG	5400
Qy	553	TTTTCGGCAGGTGTCTGGGCTGTTCGGCGCTGGCGGCAAGGCTCTTCCTTGGATCGGCAC	6122
Db	541	CTTTCGGCAGGTGTCTGGGCTGTTCGGCGCTGGCGGCA-GGTTCCTCCTCGGAGATCGGCAC	5999
Qy	613	CCCTCCCTGGGCTCATCTTAAAGCTGGCCCCCTTCCTAACCTACTCTTGGACTCTTTCAAGT	6722
Db	600	CCCTCCCTGGGCTCATCTTCAAGGCTG-CCCTTCCTACCTACTCTGGAGCTCTTCAGAGT	6588
Qy	673	TCACGTAGGGGCGTTTCGCTCCCGAG	697
Db	659	TCACGTAGGGGCGCTTCGCTTCGCCG	683

RESULT 5	LOCUS	DEFINITION	561 bp	linear	EST 04-MAY-2000
AW763237	AW763237	us70d09.v1 NCI-CGAP Mam3 Mus musculus cDNA clone IMAGE:115633 5' similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS ; mRNA sequence.			

ACCESSION	AM763237
VERSION	AM763237.1
KEYWORDS	GI:7695174
SOURCE	EST.
ORGANISM	house mouse, Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 561)
AUTHORS	NCI-CCGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Other_ESTs: ur70d09.x1

JOURNAL	Unpublished (1997)
COMMENT	Other_ESRs: ur70d09.xl
	Contact: Robert Strausberg, Ph.D.
	Email: cgapbs-remail.nih.gov
	Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
	cDNA Library Preparation: Life Technologies, Inc.
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
	DNA Sequencing by: Washington University Genome Sequencing Center
	Clone distribution: NCI-CGAP clone distribution information can be
	found through the I.M.A.G.E. Consortium/LNLN at:
	Image.lnl.gov/image/html/tiresources.shtml
	MGI:1058389
	Seq primer: -40RP from GIBCO
	High quality sequence stop: 433.
FEATURES	Location/Qualifiers
source	1..561

```
FEATURES
source
1. 561
Location/Qualifiers
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone_image="IMAGE:3155633"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
```

BASE COUNT	108 a	158 c	194 g	100 t	1 others
ORIGIN	/site=2*organ: mammary; Vector: pCMV-SPORT6; site_1: SalI; site_2: NotI; Cloned unidirectionally. Primer: oligo dt library constructed by Life Technologies. Investigators providing samples: Lothar Hennigsausen/Chu Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). "				
Query Match	43.2%	Score 50.4	DB 10	Length 561	
Best Local Similarity	99.8%	Pred. NO. 3,1e-120			
Matches 504; Conservative	0; Mismatches 1;	Indels 0;	Gaps 0;		

OY	61	GGGAGACTGGGCAACGCTGTCTGCCAGAGCCTTCTCAGAGAGCTGACACAGAGA	120
Db	117	GGGGAGCGTGGGCAACGCGTGCTGCACCAGAGCCCTTCTTCAGAGAGCGTGAAGAGA	176
OY	121	CCGGCGGGAGGCCCCCTGAACTGAATCCCACAGACAGAGAAAGCCAGATGTGTACCTTT	180
Db	177	CCGGCGGGAGGCCCCCTGAACTGAATCCCCAACAAGAGAAAGCCAGATGTGTGTACCTTT	236
OY	181	CTTGGAAACACTAGTCCGGGCTCGGAAGAAGTGCTCTTAAGAGCCGGAAGGCGCGGCTCG	240
Db	237	CTTGGAAACAACTAGTCCGGGCTCGGAAGAAGTGCTCTTAAGAGCCGGAAGGCGCGGCTCG	296
OY	241	CCGAGCTATTTCACGCCCATTTATGAGGTTTCATCCCTCGGCCAGACAGATGGAGACAAGC	300
Db	297	CCGAGCTATTTCACGCCCATTTATGAGGTTTCATCCTCGGCCAGACAGATGGAGACAAGC	356
OY	301	AGGTGTGATGAGGGAACGTAGTAGGCTGGGAAGAGAACCAAAATCAAACAGCTCCAGCCCTCT	360
Db	357	AGGTGTGATGAGGGAACGTAGTAGGCTGGGAAGAGAACCAAAATCAAACAGCTCCAGCCCTCT	416
OY	361	GCGCTACGACCGCCAGATTTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACCTGTA	420
Db	417	GCGCTACGACCGCCAGATTTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACCTGTA	476
OY	421	CTGTACAGTGACATTTGATGAGGGAAGAAAGCTGTCTACTCTGAAGCTGAGACTGTGCTGAA	480
Db	477	CTGTACAGTGACATTTGATGAGGGAAGAAAGCTGTCTACTCTGAAGCTGAGACTGTGCTGAA	536
OY	481	CGGTGTGCTGGCCCTCGGCGTGCCTTG	505
Db	537	CGGTGTGCTGGCCCTCGGCGTGCCTTG	561

RESULT 6	
LOCUS	BEG28951
DEFINITION	BEG28951 533 bp mRNA linear EST 25-AUG-2000
ACCESSION	U030C03.Y1 Soares_mammary_gland.NMIMG Mus musculus cDNA clone
VERSION	IMAGE:3373444.5 similar to TR:054907 054907 TNF-RELATED WIRAK
KEYWORDS	INDUCER OF APOPTOSIS ; , mRNA sequence.
SOURCE	BEG28951
ORGANISM	BEG28951.1 GI:9911639 EST. house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 533)
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapsr@mail.nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium ( <a href="mailto:info@image.lnl.gov">info@image.lnl.gov</a> ) for further information. <a href="mailto:MGI:1083048">MGI:1083048</a>









Db	434	TCCTCTGGGGCCCCCAGCTCCGCTCTTGCCAGAGTGTGTGGGCTGTGGCCCTGGGGCCAGG	493
OY	593	TCCTCTGGGGCCCCCAGCTCCGCTCTTGCCAGAGTGTGTGGGCTGTGGCCCTGGGGCCAGG	652
Db	494	TCCTCTGGGGCCCCCAGCTCCGCTCTTGCCAGAGTGTGTGGGCTGTGGCCCTGGGGCCAGG	553
OY	653	TACTTTGGACATCTTTCAGAGTTCACGTAGAGAGGGGCTGTGTCTCCCAATTCCTTAAACTTTC	712
Db	554	TACTTTGGACATCTTTCAGAGTTCACGTAGAGAGGGGCTGTGTCTCCCAATTCCTTAAACTTTC	613
OY	713	CTGTGCTCGACAGAGCATCACACACCTGCTCCCTACCCACCCACCTACCTTCACACCCCTC-G	771
Db	614	GGGGGCTTC-----CCTGACACAGCTCTCTGGGACACCCGCTCCCTCTGTGCCCCACCTCAG	668
OY	772	CTGTGCTCGACAGCATCACACACCTGCTCCCTACCCACCCACCTACCTTCACACCTT	829
Db	669	CCGCTCTTGTGCTCAGACCTGCTCCCTCCTTACAGAGCTGCTGGGCGCTTTCACAGTGT	728
OY	830	TTCATTC	837
Db	729	TTCATTC	736

**QUERY MATCH**  
 Best Local Similarity 40.5%; Score 473; DB 13; Length 834;  
 78.0%; Pred. No. 4,4e-112;

**BASE COUNT**  
 154 a 287 c 223 g 170 t

**ORIGIN**

**REFERENCE**  
 AUTHORS NIH-MGC  
 TITLE NIH-MGC  
 JOURNAL NIH-MGC  
 COMMENT NIH-MGC

**FEATURES**  
 source

1. 834  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5206217"  
 /clone\_lib="NIH-MGC\_122"  
 /lab\_host="DH10B"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC library."  
 High quality sequence stop: 772.  
 Location/Qualifiers

[illegible]



```

|||||
44 GAACGAATCCCGACAGAGAAAGCCAGATCTGCGCTTCTGAAACGACATAGTT 103
OY 197 CGGCTCGAAGAGTCTCTTAAGCCGAGAGCGGCTGCGGACTATTGACGC 256
Db 104 CGGCTCGAAGAGTCTCTTAAGCCGAGAGCGGCTGCGGACTATTGACGC 163
OY 257 CATATGAGTTCATCTCGGCGGAGAGAGATGAGACACAGCAGTGTGGAGCA 316
Db 164 CATATGAGTTCATCTCGGCGGAGAGAGATGAGACAGCAGTGTGGAGCA 223
OY 317 GTGAGTGGCTGGAGAGACCAAAATCAACAGCTCCAGCCCTTGGCTACGACGCCAG 376
Db 224 TTGAGTGGCTGGAGAGACCAAAATCAACAGCTCCAGCCCTTGGCTACGACGCCAG 283
OY 377 ATGAGGGAATTTACATCATCAGAGGCTGGGCTTACTACTGTACTGTAGGTGCACTTT 436
Db 284 ATGAGGGAATTTACATCATCAGAGGCTGGGCTTACTACTGTACTGTAGGTGCACTTT 343
OY 437 GATGAGGAAGAGCTGTCTACCTGAGAGCTGACTGTGTAAGAGGTGTGGGCTG 496
Db 344 GATGAGGAAGAGCTGTCTACCTGAGAGCTGACTGTGTAAGAGGTGTGGGCTG 403
OY 497 CGCTGCTGGAGAAATTTCTACAGCAGCAGCAGCAGCTTCCTGGGCGCCAGCTCGCTTTG 556
Db 404 CGCTGCTGGAGAAATTTCTACAGCAGCAGCAGCAGCTTCCTGGGCGCCAGCTCGCTTTG 463
OY 557 TGCCAGAGTCTGGGCTGTGGGCTGGGCGCCAGGCTTTCCTTGGAGTCCGACCTTC 616
Db 464 TGCCAGAGTCTGGGCTGTGGGCTGGGCGCCAGGCTTTCCTTGGAGTCCGACCTTC 523
OY 617 CCTGAGGCTCATCTTAAGAGCTGCCCCCTTCTTAACCTTGTGAGCTTTCAGTTTAC 676
Db 524 CCTGAGGCTCATCTTAAGAGCTGCCCCCTTCTTAACCTTGTGAGCTTTCAGTTTAC 583
OY 677 TGAGGGGCTTGTCTTCCAGATTTCTTAACCTTGTGAGCTTTCAGTTTAC 736
Db 584 TGAGGGGCTTGTCTTCCAGATTTCTTAACCTTGTGAGCTTTCAGTTTAC 636
OY 737 CCTCCTT---ACCCACCCCACTCCTCCAGATTTCTTAACCTTGTGAGCTTTCAGTTTAC 793
Db 637 CCTCCTTGTGGACACCGCTGCTCCCTTGTGGCCACCTCAGCAGCTTGTGCTCAGACCTGC 696
OY 794 CTCTCC 799
Db 697 CCTCCTC 702

RESULT 14
AA221610 471 bp mRNA linear EST 13-FEB-1997
LOCUS my18d09.r1 Barstead mouse heart MRLR3 Mus musculus cDNA clone
DEFINITION IMAGE:696209 5', mRNA sequence.
ACCESSION AA221610
VERSION AA221610.1 GI:1840863
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 471)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dierich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The Washu-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maira M/Mouse EST Project
Washu-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

```

```

Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (Info@image.lnl.gov) for further information.
MGI:429769
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 301.
Location/Qualifiers
FEATURES
source
1..471
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:696209"
/clone_lib="Barstead mouse heart MRLR3"
/sex="mixed"
/tissue_type="heart"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="Organ: heart; Vector: p7773D-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoRI; Site.2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACGATCTGAAGTGGAGCGGCGGCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CTTGATTCGGTACC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7773 vector.
Library constructed by Bob Barstead."
BASE COUNT 91 a 141 c 128 g 111 t
ORIGIN
Query Match 37.5%; Score 437.8; DB 9; Length 471;
Best Local Similarity 97.0%; Pred. No. 5.1e-103;
Matches 457; Conservative 0; Mismatches 12; Indels 2; Gaps 1;
OY 274 TCGGCCAGCAGAGATGAGACACAGAGTGTGATGGACGTGAGTCTGGAGAGA 333
Db 3 TCGGATCCAAAGAGATGAGACACAGAGTGTGATGGACGTGAGTCTGGAGAGA 62
OY 334 GACCAAAATCAACAGCTCCAGCCCTGCGCTTCAGACCCGACAGATTGGGAATTTACAT 393
Db 63 GACCAAAATCAACAGCTCCAGCCCTGCGCTTCAGACCCGACAGATTGGGAATTTACAT 122
OY 394 CATCAGAGGCTGGGCTTACTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGT 453
Db 123 CATCAGAGGCTGGGCTTACTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGT 182
OY 454 CTACCTGAAGCTGAGCTTGTGTGTAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 513
Db 183 CTACCTGAAGCTGAGCTTGTGTGTAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 242
OY 514 CTGAGCCACAGCAGCAAGCTCTGAGGCGCCGAGCTCCGTTGTGACAGGTGTGGGCT 573
Db 243 CTGAGCCACAGCAGCAAGCTCTGAGGCGCCGAGCTCCGTTGTGACAGGTGTGGGCT 302
OY 574 GTTGCCGCTGCGGCGCAGGCTTCTCCTTGGATCCGACACCTCCCTGGGCTCATCTTA 633
Db 303 GTTGC--GCTGGCGGAGGGCTTCTCCTTGGATCCGACACCTCCCTGGGCTCATCTTA 360
OY 634 GGTGCCCCCTTCTTAACCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCT 693
Db 361 GGTGCCCCCTTCTTAACCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCT 420
OY 694 CCGATTTCCTTAACCTTCCCTGGGCTCCAGGACATCACACACCTCCCTTA 744
Db 421 CCGATTTCCTTAACCTTCCCTGGGCTCCAGGACATCACACACCTCCCTTA 471

RESULT 15
BO671259 963 bp mRNA linear EST 15-JUL-2002
LOCUS BO671259
DEFINITION AGENCOURT_8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
ACCESSION BO671259
VERSION BO671259.1 GI:21782093

```

KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.  
REFERENCE 1 (bases 1 to 963)  
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: LNCM2456 row: 1 column: 13  
High quality sequence stop: 565.  
Location/Qualifiers  
1. 963  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="6274716"  
/clone\_lib="NIH\_MGC\_102"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pOT87; Site\_1: XhoI;  
Site\_2: EcoRI; CDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAAGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-CDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 194 a 326 c 260 g 183 t  
ORIGIN

Query Match 36.1%; Score 421.6; DB 14; Length 963;  
Best Local Similarity 78.5%; Pred. No. 1.2e-98;  
Matches 557; Conservative 0; Mismatches 144; Indels 9; Gaps 4;

```

OY 193 AGTCCGGCTCGAAGAAGTCTCTTAAGCGCGGAGCGGCTCGCCGACTATTGC 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 AGTTCGGCTCGCAGAGTGCACCTTAAGCGCGGA-AAACGGGCTCGAAGAGCATCGC 59
OY 253 AGCCATTATGAGGTTCATCTCGGCGAGAGATGAGACACAGCAGAGGTGATGG 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 60 AGCCATTATGAGTTCATCCAGCACTCGACAGCAGAGCGAGCGAGGTGACGG 119
OY 313 GACAGTGTGCTGCGAAGAACAAATCAACAGCTCCAGCCCTCGGCTAGACCG 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 120 GACAGTGTGCTGCGAAGAACAAATCAACAGCTCCAGCCCTCGGCTAGACCG 179
OY 373 CCAGATTGGGGAATTTCAGTCACTCAGGCGTCTACTACTGTAAGTCA 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 180 CCAGATTGGGGAATTTCAGTCACTCAGGCGTCTACTACTGTAAGTCA 239
OY 433 CTTTGATGAGGGAAGGCTGTCTACTGAAAGCTTGAGTGAACGTTGCTGGC 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 CTTTGATGAGGGAAGGCTGTCTACTGAAAGCTTGAGTGAACGTTGCTGGC 299
OY 493 CTTGCGCTGCTGGAATTCAGCAGACAGCAAGCTTCCTGGGCCCCAGTCCG 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 300 CTTGCGCTGCTGGAATTCAGCAGACAGCAAGCTTCCTGGGCCCCAGTCCG 359
OY 553 TTTGTGCAAGTGTCTGCGGTGTTGCCGTCAGGCGAGGCTTCCCTTGATCCGAC 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 360 CTTGTGCAAGTGTCTGCGGTGTTGCCGTCAGGCGAGGCTTCCCTTGATCCGAC 419
OY 613 CTTGCCCTGGGCTCATTTAAGGCTGCCCTTCTTAACCTATTGGACTCTTCAAGT 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 420 CTTCCCTGGGCCCCATCTAAGGCTGCCCTTCTCAGCTTCTGAGCTTCCAGGT 479
OY 673 TCACTGAGGGGCTTGTGCTCCAGATTTCCTTAACCTTCCCTGCTCCAGAGCATCAC 732
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 480 TCACTGAGGGGCTTGTGCTCCAGATTTCCTTAACCTTCCCTGCTCCAGAGCATCAC 534
OY 733 CACAGCTCCCTACCCACACCCGCTCCAGCCGCTC-GCTGCTCTGCTGCTGCT 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 535 CAGCTCTGTGGGACCCGCTCCCTGCTCCAGCCGCTCTGCTCTGCTCCAGACT 594
OY 792 GTCTCTCC--TCAAGGAGCAGACGCTGTTCATGTTTCCATTGACAGAGCTATCC 849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 595 GCCCTCTCTTAAGGCTGCTGCGGCTGTTCAAGTGTTCATGTTTCCATTGACAGAG 654
OY 850 TTGCTCTTAAACATCCATCCACCAACAATTCACCTCACTAGCTC 899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 655 GTATCCCACTTATCTTACAACTCCCCCAAGGCGCACTCTGCAGCTC 704

```

Search completed: March 31, 2003, 11:47:15  
Job time : 2156.97 secs

1	1.168	100.0	1.168	19	AAV18259	Mus musculus tumou
2	699.4	59.9	701	20	Mouse TNRJ3 DNA.	
3	628.6	53.8	1353	21	AAAX23425	Human TNRJ3 DNA.
4	628.6	53.8	1353	24	AAAX9717	Human PRO207 CGNA
5	628.6	53.8	1421	20	ABK40255	CGNA encoding huma
6	618.2	52.9	1364	24	AAAY56000	Human tumour necro
7	614.6	52.6	1373	19	ABK34881	Human cDNA encodin
8	597.8	51.2	1236	19	AAV18600	Human Spleens tumou
9	597.8	51.2	1236	19	AAV47613	TNF related endothe
				22	AAAD04350	Human TREPA (TNF r

10	532.8	44.8	1030	20	AA5233424	Human TRRL3 DNA.
11	498.8	42.7	1898	22	AA5230964	Expression vector
12	82.8	7.1	282	16	AA7221990	human gene signal
13	69.8	6.0	195	24	ABR29540	Colon adenocarcinoma
14	65	5.6	65	24	ABN55975	Mouse spliced trans
15	46.6	4.0	5352	24	AA230228	Human PKD1 gene.
16	46.6	4.0	5352	19	AA794101	Human PKD1 gene.
17	46.6	4.0	5357	17	AA716551	Human polydystic k
18	46.6	4.0	5357	19	AA794107	Human PKD1 locus b
19	45.8	3.9	105355	24	ABR94407	DNA encoding endob
20	45	3.8	1337	20	AA217263	Human gene express
21	44.8	3.8	1000	21	AAA02484	Human colon cancer
22	43.6	3.7	16167	24	ABL70254	Chemically treated
23	43.6	3.7	16167	24	ABL33083	Human immune syste
24	43.6	3.7	16167	24	ABL34529	Human metastasis a
25	43.2	3.7	324	24	AAH51826	Human CDNA clone (
26	43.2	3.7	720	22	AAH05001	Human CDNA sequenc
27	43.2	3.7	2260	22	AAH18456	Human DNA for a no
28	43.2	3.7	2272	22	AA534689	Human DNA for a no
29	43.2	3.7	2272	22	AA534690	Human DNA for a no
30	43.2	3.7	8604	24	ABO71065	Listeria monocytog
31	43	3.7	524	24	ABO34340	Oligonucleotide fo
32	43	3.7	524	24	ABO34341	Oligonucleotide fo
33	42.8	3.7	1166	20	AAZ19440	Oligonucleotide fo
34	42.8	3.7	1166	20	AAZ19228	M. tuberculosis re
35	42.6	3.6	35	20	AAA31883	M. tuberculosis re
36	42.6	3.6	556	24	ABO36594	Primer used in the
37	42.6	3.6	556	24	ABO36595	Oligonucleotide fo
38	42.6	3.6	676	24	ABO36768	Oligonucleotide fo
39	42.6	3.6	676	24	ABO36769	Oligonucleotide fo
40	42.4	3.6	1286	24	AB199656	Mouse ischaemic co
41	42.4	3.6	10732	21	AAH10594	gene encoding a su
42	41.8	3.6	14006	24	AB133959	Human immune syste
43	41.6	3.6	1164	24	ABO68833	Listeria monocytog
44	41.4	3.5	320	21	AAZ18186	Primer used in the
45	41.4	3.5	1235	20	AAZ16147	Human gene express

PA (UYE-) UNIV GENEVA FACULTY MEDICINE.  
 XX Browning JL, Chicheportiche Y;  
 XX  
 DR MPI: 1998-145619/13.  
 DR P-PSDB: AAM47524.  
 XX  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 2: Pages 45-46; 69pp; English.  
 XX  
 CC The sequence is that encoding mouse tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments to be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. The DNA sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of prepare probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 CC  
 XX  
 SQ Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;  
 Query Match 100.0%; Score 1168; DB 19; Length 1168;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-307;  
 Matches 1168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GCCCAGCTCCGTTTGTGCGAGGTCTGGGCTGTGTCGCGGCGGAGGCTTTCCT 600  
 |||||||  
 Db 541 GCCCAGCTCCGTTTGTGCGAGGTCTGGGCTGTGTCGCGGCGGAGGCTTTCCT 600  
 QY 601 TCGGATCCGACACCTCCCTGGGCTCATCTTAAAGCTGCCCTTCTTAACACTTTGG 660  
 |||||||  
 Db 601 TCGGATCCGACACCTCCCTGGGCTCATCTTAAAGCTGCCCTTCTTAACACTTTGG 660  
 QY 661 ACTCTTCAAGTTCACAGAGGGGCTTGGCTCCAGATTCTTAAACTTCCCTGGCTC 720  
 |||||||  
 Db 661 ACTCTTCAAGTTCACAGAGGGGCTTGGCTCCAGATTCTTAAACTTCCCTGGCTC 720  
 QY 721 CAGAGCATCAGACACACCTCCCTACCCACCCACCTCTCCACCCCTGCTGCTCTT 780  
 |||||||  
 Db 721 CAGAGCATCAGACACACCTCCCTACCCACCCACCTCTCCACCCCTGCTGCTCTT 780  
 QY 781 GGTCCAGTCTGCTCTCTCTCAAGGACGACGACTTGTTCACAGTTCATTCACCA 840  
 |||||||  
 Db 781 GGTCCAGTCTGCTCTCTCTCAAGGACGACGACTTGTTCACAGTTCATTCACCA 840  
 QY 841 GAGTATCCCTTGTCTCTTAACATCCCATCCGACCACTATCCACCTCAGAGCTCC 900  
 |||||||  
 Db 841 GAGTATCCCTTGTCTCTTAACATCCCATCCGACCACTATCCACCTCAGAGCTCC 900  
 QY 901 CAAAGCCCTACTTATCCCTGACTCCGCCACCCACTCACCAGACAGCTTTATTGACT 960  
 |||||||  
 Db 901 CAAAGCCCTACTTATCCCTGACTCCGCCACCCACTCACCAGACAGCTTTATTGACT 960  
 QY 961 TTGTGACACGACGACATGAGTGGCTGAGCTGTGGGACGAGGACGACAGACCTGGGAC 1020  
 |||||||  
 Db 961 TTGTGACACGACGACATGAGTGGCTGAGCTGTGGGACGAGGACGACAGACCTGGGAC 1020  
 QY 1021 TAGGCCAAGATTGCCAAGCTGTGAGGGGAGAGAGCTGGGAGACGCTCCCTGGATCC 1080  
 |||||||  
 Db 1021 TAGGCCAAGATTGCCAAGCTGTGAGGGGAGAGAGCTGGGAGACGCTCCCTGGATCC 1080  
 QY 1081 CTGTGATTTTGAAGAAATCTATTTTATTTATTTGTGACAAATGTTAAATGATAT 1140  
 |||||||  
 Db 1081 CTGTGATTTTGAAGAAATCTATTTTATTTATTTGTGACAAATGTTAAATGATAT 1140  
 QY 1141 TAAAGAGAAATTAATCATGATTTCTCTTC 1168  
 |||||||  
 Db 1141 TAAAGAGAAATTAATCATGATTTCTCTTC 1168

RESULT 2  
 AAX23425  
 ID AAX23425 standard; DNA: 701 BP.  
 AC AAX23425;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Mouse TNRL3 DN.  
 XX  
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse; ss.  
 XX  
 OS Mus sp.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1..636  
 FT /\*tag= a  
 FT /\*product= "TNRL3"  
 PN W09911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX



PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIV ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 DR P-PSDB: AAW93591.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Example VII; Fig 13B; 156pp; English.

CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: ApO4, ApO6, ApO8 and ApO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. ApO4 is useful for diagnosing prostate cancer  
 CC by determining levels of ApO4 in an individual. Prostate cancer can also  
 CC be treated using ApO4 selective binding agents linked to a therapeutic  
 CC moiety. ApO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. ApO4 polypeptides/ active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in ApO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using ApO4 polypeptides/active  
 CC fragments and ApO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4  
 CC activity. The method is performed in vivo or in vitro. ApO polypeptides  
 CC are all useful as immunogens for preparing antibodies. ApO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. ApO8 was transfected to human breast carcinoma cell line  
 CC MCP-7, and induced apoptosis.

CC  
 XX  
 XX Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Query Match 59.9%; Score 699.4; DB 20; Length 701;  
 Best Local Similarity 99.9%; Pred. No. 1.1e-179;

Matches 700; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 44 CTGGTCGTCGTCAGCCCTGGGGAGCTGGGCAACGCTGTCGCCACAGAGCTTCAGAG 103  
 DB 1 CTGGTCGTCGTCAGCCCTGGGGAGCTGGGCAACGCTGTCGCCACAGAGCTTCAGAG 60  
 OY 104 GAGCTGACAGAGAGAGACCGCGGAGCCCTGTAAGTAATCCCAAGAGAGAAAGC 163  
 DB 61 GAGCTGACAGAGAGAGACCGCGGAGCCCTGTAAGTAATCCCAAGAGAGAAAGC 120  
 OY 164 CAGAGTGTGTACTCTTTTGTGAACACTAGTCGCGCTCGAAGAGTCTCTTAAAGGC 223  
 DB 121 CAGAGTGTGTACTCTTTTGTGAACACTAGTCGCGCTCGAAGAGTCTCTTAAAGGC 180  
 OY 224 CGGAAGGGGGGCTCGCGGAGCTATTGACAGCCCATTTAGAGTTCATCTCGGCACAGA 283  
 DB 181 CGGAAGGGGGGCTCGCGGAGCTATTGACAGCCCATTTAGAGTTCATCTCGGCACAGA 240  
 OY 284 CAGATGAGACACAAGAGAGTGTGGAGAGTGGAGTGGGAGAGAGCAAAATTC 343  
 DB 241 CAGATGAGACACAAGAGAGTGTGGAGAGTGGAGTGGGAGAGAGCAAAATTC 300  
 OY 344 AACAGCTCCAGCCCTCTGCGCTACAGCCGAGATTGGGGAATTTACAGTCATCAGGGCT 403  
 DB 301 AACAGCTCCAGCCCTCTGCGCTACAGCCGAGATTGGGGAATTTACAGTCATCAGGGCT 360  
 OY 404 GGGCTCTACTACTCTACTCTGAGTGCATTTGATGAGGAGAAAGCTGTCTACTGAG 463  
 DB 361 GGGCTCTACTACTCTACTCTGAGTGCATTTGATGAGGAGAAAGCTGTCTACTGAG 420

OY 464 CTGACCTTGCTGTGTAAGAGGTGTGCGGCCCTGCGCTGTGGAAGAAATTCACGCACA 523  
 DB 421 CTGACCTTGCTGTGTAAGAGGTGTGCGGCCCTGCGCTGTGGAAGAAATTCACGCACA 480  
 OY 524 GCAGCAAGCTCTCTGTGGGCCCCAGCTCGTTTGTGCAAGAGTGTGCGCTGTGCGCTG 583  
 DB 481 GCAGCAAGCTCTCTGTGGGCCCCAGCTCGTTTGTGCAAGAGTGTGCGCTGTGCGCTG 540  
 OY 584 CGGCGAGGGTCTTCCCTTCGATTCGCAACCCCTCCCTGGGCTCATCTTAAGGCTCCCCC 643  
 DB 541 CGGCGAGGGTCTTCCCTTCGATTCGCAACCCCTCCCTGGGCTCATCTTAAGGCTCCCCC 600  
 OY 644 TTCTTAACCTACTTTGAGACTCTTTCAGATTCAGTAGGGGCTTGTCTCCAGATTCT 703  
 DB 601 TTCTTAACCTACTTTGAGACTCTTTCAGATTCAGTAGGGGCTTGTCTCTCCAGATTCT 660  
 OY 704 TAACTTTCCTGCTCCAGAGACATACACACCTCTCCCTA 744  
 DB 661 TAACTTTCCTGCTCCAGAGACATACACACCTCTCCCTA 701

RESULT 3  
 AAA49717  
 ID AAA49717 standard; cDNA: 1353 BP.  
 XX  
 XX AAA49717;  
 AC  
 AC 25-SEP-2000 (first entry)  
 DT  
 XX  
 XX Human PRO207 cDNA clone DNA30879-1152.  
 DE  
 XX PRO207; human; antitumor; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 58..807  
 FT sig\_peptide /\*tag= a  
 FT mat\_peptide /\*tag= b  
 FT 178..804  
 FT /\*tag= c  
 FT  
 PN WO200037638-A2.  
 XX  
 XX 29-JUN-2000.  
 PD  
 XX  
 XX 02-DEC-1999; 99WO-US28565.  
 PE  
 XX 22-DEC-1998; 98US-0113296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX  
 XX (GENE) GENENTECH INC.  
 PA  
 XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Piltl RM, Wood WI;  
 DR WPI: 2000-442668/38.  
 DR P-PSDB: AAY95338.  
 XX  
 XX Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or

PT PRO866 -  
 XX  
 PS Claim 20: Fig 3; 172pp; English.  
 XX  
 CC The present sequence is that of cDNA clone DNA30879-1152  
 CC (ATCC 209338) encoding human PRO207 (see AAY95338), which shows  
 CC homology to several members of the tumour necrosis factor family,  
 CC especially human lymphotoxin (23.48). The cDNA was identified in a  
 CC foetal kidney cDNA library following identification of an expressed  
 CC sequence tag with homology to human Apo-2 ligand. A claimed method  
 CC for inhibiting the growth of a tumour cell comprises exposing the  
 CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO224, PRO328,  
 CC PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see  
 CC AAY95337-49), their agonists or chimeric polypeptides incorporating  
 CC them. The tumour is especially a cancer selected from breast,  
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and  
 CC central nervous system cancer, melanoma and leukaemia. Nucleic  
 CC acids encoding PRO179 etc. are used in the recombinant production  
 CC of the antitumour polypeptides.  
 XX  
 SO Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;  
 Query Match 53.8%; Score 628.6; DB 21; Length 1353;  
 Best Local Similarity 76.2%; Pred. No. 2,5e-160;  
 Matches 933; Conservative 0; Mismatches 219; Indels 73; Gaps 10;

QY 722 AGGAGCATACACACCTCCCTACCCACCACCTCTCTCAACCCCTC-6GTGCTCTT 780  
 DB 850 -----CTGCACAGCTCTCTGGGACCCGGTCCCTCTGCCCCACCTCAGCCGCTT 904  
 QY 781 GGTCCAGTCCCTGTCTCC-TCMAAGCAGCCAGAGCTTTGTACATGTTCCATTC- 837  
 DB 905 GCTCCAGACTGCCCCCTTCTAGAGGCTGCTGGGCTGTTCACGTGTTTCCATCCC 964  
 QY 838 -----ACAGCTATTCCTTGTCTTACATCCCATCCACCACTATCCACCTC 891  
 DB 965 ACATAATATAGATTTCCACACTCTTATCTTACACCTCCACAGCCACTCCACCTC 1024  
 QY 892 ACTAGCTCCCAAGCCCTCA-----TTATCCCTGACTCCCAACCACT 936  
 DB 1025 ACTAGCTCCCAATCCCTCAACCTTTGAGGCCCCCAAGTATCTGCACTCCCTGCGCA 1084  
 QY 937 CACCCGACACAGTGTATTTAGACTTTGTGAC----- 968  
 DB 1085 CAGACCCCGCATTTGTTCACCTCTGTGCGCAAGATGGTCCAGAGACCC 1144  
 QY 969 -----CAGGACTGATATGGCTGTGACCTGTGCGAGGAAGCCAGAACCTGGACTAG 1023  
 DB 1145 CACTTACGACCTAAAGAGGGGCTGACCTGGCGGAGGAAGCAAGACTGGCTAG 1204  
 DB 1205 GCCAGAGGTTCCCAATGTAGGGGGAGG-AMCAGACAGCTCCTCCCTGAGAAATTC 1263  
 QY 1024 GCCAGAGTTCCTCACTGTGTAGGGGAGAGAGCTGGGAGCAAGCTCCTCGA----TC 1079  
 DB 1205 GCCAGAGGTTCCCAATGTAGGGGGAGG-AMCAGACAGCTCCTCCCTGAGAAATTC 1263  
 QY 1080 CCTGTGATTTTCAAA--AGATCTATTTTATTTATTTGTACAAATGT---TAAAT 1134  
 DB 1264 CCTGTGATTTTCAAA--AGATCTATTTTATTTATTTATTTGTACAAATGT---TAAAT 1134  
 QY 1135 GATATTTAAAGATTAATCATGA 1159  
 DB 1324 GATATTTAAATAGATTAATCATGA 1348

RESULT 4  
 ABK40255  
 ID ABK40255 standard; cDNA: 1353 BP.  
 XX  
 AC ABK40255;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE cDNA encoding human PRO207 polypeptide.  
 XX  
 KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;  
 KW Leukemia; neuronal disorder; stromal disorder; blastocellular disorder;  
 KW Inflammatory disorder; immune disorder; angiolethic disorder;  
 KW gene therapy; cytostatic; neuroprotective; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200153486-A1.  
 PD 26-JUL-2001.  
 PF 11-FEB-2000; 2000MO-US03565.  
 XX  
 PR 08-MAR-1999; 99MO-US05028.  
 PR 11-MAR-1999; 99MO-US05028.  
 PR 11-MAY-1999; 99MO-US05028.  
 PR 02-JUN-1999; 99MO-US12252.  
 PR 22-JUN-1999; 99MO-US140650P.  
 PR 22-JUN-1999; 99MO-US140653P.  
 PR 20-JUL-1999; 99MO-US144758P.  
 PR 26-JUL-1999; 99MO-US146222P.  
 PR 28-JUL-1999; 99MO-US146222P.  
 PR 31-AUG-1999; 99MO-US149395P.  
 PR 31-AUG-1999; 99MO-US15689P.  
 PR 01-SEP-1999; 99MO-US20111.









PN W09835061-A2.  
 PD 13-AUG-1998.  
 XX 12-FEB-1998; 98WO-US02859.  
 XX 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX (ABBO ) ABBOTT LAB.  
 PA  
 XX Wiley SR;  
 PI WPI: 1998-447255/38.  
 DR P-PEDB; AAM29745.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 PS Claim 11; Page 123-4; 142pp: English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:  
 SQ  
 Query Match 51.2%; Score 597.8; DB 19; Length 1236;  
 Best Local Similarity 75.2%; Pred. No. 5,6e-152;  
 Matches 879; Conservative 0; Mismatches 222; Indels 68; Gaps 8;

Db 493 TGTCAAGTGCACCTTGATGAGGGAAGGCTGTCTACTACCTGAGACTTGTGATGAT 552  
 Qy 482 GGTGTGCTGGCCCTGCGCTGCTGGAGAATTCTACGACAGCAAGCTCTCTGGG 541  
 Db 553 GGTGTGCTGGCCCTGCGCTGCTGGAGAATTCTACGACAGCAAGCTCTCTGGG 612  
 Qy 542 CCCAGCTCCTGTTGGCAGGAGTGTGGGCTGTTGGCCGCTGGGCAAGGCTTCCCTT 601  
 Db 613 CCCAGCTCCTGTTGGCAGGAGTGTGGGCTGTTGGCCGCTGGGCAAGGCTTCCCTT 672  
 Qy 602 CGGATCCGACCCCTCCCTGGGCTCATTTAAGGCTGCCCCCTCTCACTACTTTGGA 661  
 Db 673 CGGATCCGACCCCTCCCTGGGCTCATTTAAGGCTGCCCCCTCTCACTACTTTGGA 732  
 Qy 662 CTCTTCAAGTTCACTGAGGGGCTTGTCTCCAGATTCTTAACTTTCCCTGGCTCC 721  
 Db 733 CTCTTCAAGTTCACTGAGGGGCTTGTCTCCAGATTCTTAACTTTCCCTGGCTCC 792  
 Qy 722 AGGAGCATCACACACCTCTACCCACCCGCTCTCAACCCCTC-GGTGGCTCTT 780  
 Db 793 -----CTTGACAGCTCTTGGGACCCGGTCCCTCTGCCCCACCTTACGCCGCTTT 847  
 Qy 781 GGTCCAGTCTGTCTCTCC-TCAAAGGAGCCAGAGCTGTTACATGTTCCATTCG- 837  
 Db 848 GCTCCAGACCTGCCCCCTCTAGAGGCTGCTGGGCTGTTACAGTGTTCATTCGCC 907  
 Qy 838 -----ACGAGCTATCTTGTCTCTTAAATCCATCCACGCAACTATCCAGCTC 891  
 Db 908 ACATAAATGACATATTTCCACTCTTATCTTACAAACCCACCGCCACTCTCCACCTC 967  
 Qy 892 ACTAGTCCCAAGCCCTTAC-----TTATCCCTGACTCCGACCCACT 936  
 Db 968 ACTAGTCCCAAGCCCTTAC-----TTATCCCTGACTCCGACCCACT 1027  
 Qy 937 CACCCGACCAAGCTTTATGACTTGTGAC----- 968  
 Db 1028 CAGACCCCGAGGCAATTTGTTTCACTGACTGTGCGCAAGAGTGGTCCAGAACCC 1087  
 Qy 969 -----CAGGACATGAGATGGGCTGGACCTGGGCGAGCAAGCAAGCAAGCTGAG 1023  
 Db 1088 CACTTGAAGCACTAACAGGGGCTGACCTGGGCGAGCAAGCAAGCAAGCTGAG 1147  
 Qy 1024 GCCAGAGTTCCTCACTGAGGGGAGAGAGCTGGGCAAGCTCTCCCTGGA----TC 1079  
 Db 1148 GCCAGAGTTCCTCACTGAGGGGAGAGAGCTGGGCAAGCTCTCCCTGGA----TC 1206  
 Qy 1080 CCTGTGATTTGAAAGATPACTATTTT 1108  
 Db 1207 CCTGTGATTTTAAACAGATATTTT 1235  
 RESULT 9  
 AAD04350 standard; cDNA; 1236 BP.  
 ID AAD04350  
 XX AAD04350:  
 AC  
 XX  
 DE 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent) cDNA.  
 KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vulnery; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..750  
 FT /\*tag= a  
 FT /product= "Human TREPA (TNF related endothelium  
 FT proliferative agent)"  
 FT  
 XX

```

PN   US6207642-B1.
PD   27-MAR-2001.
PE   26-JUN-1998;    98US-0105343.
PR   12-FEB-1997;    97US-0798692.
PR   10-FEB-1998;    98US-0021705.
XX   (ABBO ) ABBOTT LAB.
PI   WILEY SR;
XX   MPI; 2001-280760/29.
DR   P-PsDB: AA000891.
PT   Inducing angiogenesis in mammal at desired sites for promoting wound
PT   healing, by administering soluble fragment of extracellular domain of
PT   tumor necrosis factor related endothelium proliferative agent protein
PT   -
XX   Example 2; Column 73-74; 53pp; English.
PS
CC   The present invention relates to extracellular signal molecules,
CC   particularly members of tumour necrosis factor (TNF) family molecules
CC   designated as TRRPA (TNF related endothelial proliferative agent).
CC   Soluble biologically active TRRPA are used to treat TRRPA-associated
CC   diseases, tumours or metastases. TRRPA is used for inducing angiogenesis
CC   in human for promoting wound healing and for vascularising grafted
CC   tissue for successful grafting and to promote tissue grafts.
CC   The present sequence is a cDNA clone ID #69050 encoding human TRRPA.
XX   Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;
SQ
Query Match          51.2%; Score 597.8; DB 22; Length 1236;
Best Local Similarity 75.2%; Pred. No. 5,6e-152;
Matches 879; Conservative 0; Mismatches 222; Indels 68; Gaps
OY   2 GTGTGAGCCTGGCGCTTGCGCTGCCTTGGCCCTTCGTGCTGTCACGCTG 61
DB   73 GCGCTGGGCGCTGGGCGCTGCGCTGCGCTGCCCTCCCTGCTGCGCGTCAGTTTG 132
OY   62 GGGAGCTGGGGCAACCGTGTCTGCCCCAGAGACCTTCTCAGAGAGAGCTGACAGAGAGAC 121
DB   133 GGGAGCGGGGGCATCGCTGTGCGCCCAGAGACCTGCCAGAGAGAGCTGTGGCAGAGAG 192
OY   122 CCCCGGAGACCCCTCGAATCGAATCCCCAGAGAGAGAAAGCCAGAGATGTGTACTTTC 181
DB   193 GACCAGAGACCCGTGGGAATCGAATCCCGACAGAGAAGAAAGCAGGATCTCGGCTTTC 252
OY   182 TTGGAACAACACTAGTCCGGCTCGAGAGAGTAGTCTCTTAAGGCCGGAAGCGCGGCTCG 241
DB   253 CTGAACCCGACTGTGTTGCGGCTCGAATAACCTGACACTTAAAGGCCGGAACACCGGCTCA 312
OY   242 CGAGCATATTGCAAGCCATTATGAGGTTCATCCGCGCCAGAGAGAGATGGAGGACAAGA 301
DB   313 AGAGCGATGCGACGCCATTATATGAGTTCATCCACAGACTCGAGACGAGACGGACGGCA 372
OY   302 GGTTGTGATGGAGCACTGAGTGGCTGGAAGAGACCAAAATCAACAGCTCCAGCCCTTG 361
DB   373 GGTGTGTGAGAGGGACAGTAGTGGCTGGGAGAGAACCAAGATCAACAGCTCCAGCCCTTG 432
OY   362 GCCTAGAGACCGCAGATTGGGGAATTTACAGTCATCAGAGGCTGGGCTTACTACTCTGTAC 421
DB   433 GCCTACAAACCGCAGATGCGGGAGTTTAATGTCACCGGGCTGGGCTTACTACTCTGTAC 492
OY   422 TTCTCAGTGTACATTGTGATAGGGAAGAGCGCTGTACTTAACTGAGCTGACTTGTGTAAC 481
DB   493 TTCTCAGTGTACATTGTGATAGGGAAGAGCGCTGTACTTAACTGAGCTGACTTGTGTAAC 552
OY   482 GGTGTGTGCGCCCTGCGCTGCTGGAAGAATTTCAAGCACAGACGAAAGCTCTCTGGG 541
DB   553 GGTGTGTGCGCCCTGCGCTGCTGGAAGAATTTCAAGCACAGACGCGGAGTTCTCTCTGGG 612

```

QY	542	CCGAGCTCGGTTTgagcagagGgTGTGGGGgTgTggccGgTccGGcCgAGggGCTTCCCTT	601
Db	613	CCCCAGCTCCGGCCTTgCCAGGgTgTGGgGCTgTggGCTTgGGCCTTgCGGcCAGgGCTCTCCCTG	672
QY	602	CGATATCCGcCACCCTCCCTGGGCTCATCTTAAAGGCTGGCCCCCTTCTTAACCTACTTTGGA	661
Db	673	CGATATCCGcCACCCTCCCTGGGCTCATCTTAAAGGCTGGCCCCCTTCTTAACCTACTTTGGA	732
QY	662	CTCTTTCAAGTTCACGTAGAGGGGCTTgCTCTCCcAGATTTCTTAAACTTTCCCTGgCTCC	721
Db	733	CTCTTTCAAGTTCACGTAGAGGGGCTTgCTCTCCcAGATTTCTTAAACTTTCCCTGgCTCC	792
QY	722	AGGAGCATTCACACACCTCCCTACCCcACCCcCAGCTTCGACcCCGCTC- GGTGGCTCTT	780
Db	793	-----CCTCGACAGCTCTCTGGGcACCCGGTCCCTCTCTGcCCcACCTCTAGcCGCTCTTT	847
QY	781	GGTCACAGTCTGTCTCTCC--TCAAAGGcAGCCAGAGCTTgTTCACATgTgTTCATTTCC-	837
Db	848	GCTCCAGACCTGGCCCTCCCTCTTAgAGGGCTGGCTGGGCTGTTTTCAGTgTTCATTTCC	907
QY	838	-----ACAGACGTATCTTCTCTCTTTCATACATCCCATCCcACcACAGATTCACCTC	891
Db	908	ACATAAATTAACAGTATTTCCcACCTCTATCTTACAAcAACCCcACCGCACCTCCACCTC	967
QY	892	ACTATGCTCCCAAAGCCCCCTAC-----TTATCCCTGACTCGCCcACCCACCT	936
Db	968	ACTATGCTCCCAAATCCCTGACCTTTTAgGGCCCCcCAGTATCTCGACCTCCCCCTGGCCA	1027
QY	937	CACCCGACCAcGgTgTgTATTTGACTTTTGTGCAC-----	968
Db	1028	CAGACCCCCcAGGGGATgTGTCTACGTACTGTgTGGcCAAGATgTGGTCCAGAAACCC	1087
QY	969	-----CAGcCAGCTGAGATGGGCTGgACCTGTgTGGcCAGAAcCCAGAAcCTGGGACTAG	1023
Db	1088	CACCTTcAGGcACATPAAGAGGGGCTGgACCTGTgTGGcCAGAAcCCAGAAcCTGGGACTAG	1147
QY	1024	GCCAGAGTTCcCAACTGCTGAGGGGGAAGAGCTGGGGACAAcCTCCCTGGA----TC	1079
Db	1148	GCCAGAGTTCcCAAAATGTAGGGGGGAGAc-AACAAGACAAcCTCTCTCCCTTGAATTC	1206
QY	1080	CCTGTGGATTTTGAAGAAATACTATTTTTT	1108
Db	1207	CCTGTGGATTTTTPAAACAGATATATATT	1235
RESULT 10			
AAx23424			
AAx23424 standard; DNA: 1030 BP.			
AAx23424:			
18-JUN-1999 (first entry)			
Human TNRL3 DNA.			
Tumour necrosis factor receptor; signal transducer molecule; TNF; ApO4; development abnormality; gestational abnormality; prostate cancer; ApO6; ApO8; ApO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease; cytoplasmic domain; immunogen; antibody preparation; breast carcinoma; apoptosis; human; ss.			
Homo sapiens.			
Key Location/Qualifiers			
FT 1..627			
FT /*tag= a			
FT /product= "TNRL3"			
.MO9911791-A2.			
11-MAR-1999.			



PE 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIV ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 XX WPI; 1999-205191/17.  
 DR P-PSDB; AAM93590.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Example VII; Fig 13A; 156pp; English.

This invention describes isolated Tumor Necrosis Factor (TNF) family receptor polypeptides: ApO4, ApO5, ApO8 and ApO9 or their active fragments and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or their active fragments. ApO4 is useful for diagnosing prostate cancer by determining levels of ApO4 in an individual. Prostate cancer can also be treated using ApO4 selective binding agents linked to a therapeutic moiety. ApO4 polypeptides are also useful for identifying selective binding agents, useful in diagnosis/treatment of disease by binding of agents to the polypeptide/active fragment which is extracellular, or expressed on the cell surface. The binding is preferably performed *in vivo*. ApO4 polypeptides/active fragments are also useful for screening for agonists and antagonists by binding and observing the change in ApO4 activity. Effective pharmacological agents useful in diagnosis or treatment of disease are also identified using ApO4 polypeptides/active fragments and ApO4 signal transducer molecules that specifically interact with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4 activity. The method is performed *in vivo* or *in vitro*. ApO4 polypeptides are all useful as immunogens for preparing antibodies. ApO4 is also useful for diagnosis/treatment of developmental or gestational abnormalities. ApO8 was transfectected to human breast carcinoma cell line MCF-7, and induced apoptosis.

SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Query Match	44.8%	Score 522.8	DB 20	Length 1030
Best Local Similarity	81.6%	Pred. No. 1.2e131		
Matches 643	Conservative	0	Mismatches 137	Indels 8
				Gaps 3

QY	53	GTACAGCTGGGGAGACTGTGGGACAGCTCTGTGCCAGGAGAGCTTCTCAGAGAGAGCTGACA	112
Db	1	GTCAAGTTTGGGAGACCCGGGACATGCTGTCCGGCCAGGAGGCTGCCCAGGAGAGAGCTGGTG	60
QY	113	GCAAGAGACCGCCGGGAGACCCCTCGAATCAATCCCGACAGAGAGAAACCCAGAGATGTG	172
Db	61	GCAAGAGAGAGACAGAGAACCCCTCGGAATCTGATCCCGACAGAGAGAAAGCCAGAGATCT	120
QY	173	GTACCTTTTCTTGGACAACAATAATGTCGGGCTCGAAGAAAGTCTCTCTTAAGGCCGGAAAGCG	232
Db	121	GGCGCTTTCCTGAACCCCACTAGTTGGGCTCTGCGAGAGTGACACTTAAGGCCGGAATAACA	180
QY	233	CGGCGTCGGCGAGCTATTTGCAAGCCCAATTAATGAGTTTACTCTGGGCCAGGACACAGATGGA	292
Db	181	CGGGCTCGAAGAGGCGATTCGCGCCCAATTAATGAATTTATCTCAGACCTTGGACAGGAGGGA	240
QY	293	GCACAAACAGAGTGTGATGTGGAGACAGTGAATGAGCTGGGAGAAAGACCAAAATCAACAGCTCC	352
Db	241	GGCAGGCGAGGTGTGTGACACGGGACACAGTGAATGAGTGGCTGGGAGGAGACCAAAATCAACAGCTCC	300
QY	353	AGCCCTGTGGCTACGACACCGCCAGATTGGGGAATTTACAGTCAATCAGAGGCTGGGCTCTTAC	412
Db	301	AGCCCTGTGGCTACTACCGCCAGATGGGAGAGTTTATAGTACACCGCGGCTGGGCTCTAC	360
QY	413	TACCTGTACTGTCAAGGTGACATTTGATGAGAGGAAAGGCTGTCTACTCGAAGCTGAGACTTG	472
Db	361	TACCTGTACTGTCAAGGTGACATTTGATGAGAGGAAAGGCTGTCTACTCGAAGCTGAGACTTG	420

QY	473	CTGGGAACGGGTGTGCTGGGCCCTGGCGTCCCTGGAAAGATTTCTCAGCCACAGCAGCAAGC	5322
Db	421	CTGGGAGTGGTGTGGCTGGGCCCTCGCGTCCCTGGAGGAATTTCTAGCCACAGCGGGCAGT	480
QY	533	TCTCTCTGGAGCCCGACAGCTCCCTTTGTGCGAGGTGTCTGGGGCTGTTGCGCGTGGCGGCCAGGG	5922
Db	481	TCCCTCGGGGGCCCCAGACTCCCGCTCTGCCAGGTGTCTGGGCTGTTGGGCTCTGGCGGCCAGGG	5400
QY	593	TCTTCCCTTCGGATCCGACACCTCCCTGGGGCTCATCTTAAGGCTGCCCCCTTCTTAAC	6522
Db	541	TCTCTCCCGTCCGATCCGACACCTCCCTGGGGCCCATCTCAAGGCTGCCCCCTTCTTACC	600
QY	653	TACTTTGGACACTTTTCAAGTTCACTAGAGGGGCTTGTCTCTCCCAAGATTTCTTAACATTTC	7122
Db	601	TACTTCGGACACTTTTCAAGTTCACTAGAGGGGCTTGTCTCTCCCAAGTGTCTCCAGAGCT	660
QY	713	CTGTGCTCAGAGAGATCACACACTCTCCCTAACCCCAACCCCACTCTCCACCCCTC-G	7722
Db	661	GCCGGCTCC-----CTTGCACAGCTCTCTGTGGGACACCGGGTCCCTCTTGCACCCACTCTCAG	715
QY	772	CTGTGCTCTGTGTCAGTCTGTCTCTCC--TCAAAAGCAGGCAGAGCTTGTTCATCATGTT	8252
Db	716	CCGGCTCTTGTGTCAGAGACTCTCCCTCCCTCTAGAGGCTGCTGGGCTGTTCACAGTGT	775
QY	830	TCCATTCG 837	
Db	776	TTCATTC 783	

RESULT 11  
AAS03964

AC AAS03964

DT 26-SEP-2001 (first entry)

DE Expression vector pDC409-L2-TWEAK fusion protein-encoding DNA

KM TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
KM rubellosis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
KM corneal graft neovascularisation; psoriasis; metastatic condition;  
KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
KM preneoplastic condition; myocardial angiogenesis; wound granulation;  
KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
KM peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
KM fusion protein.

```

05 Homo sapiens.
05 Synthetic.

```

	Key	Location/Qualifiers
FH	CDS	52..873
FT		

FT	/product=	"Fusion protein comprising a growth hormone
FT		leader, a leucine zipper multimerisation
FT		domain, and human TWEAK extracellular
FT		domain"

PN W0200145730-A2

PD 28-JUN-2001

PF 19-DEC-2000; 2000WO-US34755.

PR 20-DEC-1999; 99US-0172878.  
PR 10-MAY-2000; 2000US-0203347.

PA (IMMV) IMMUNEX CORP.

PI Wiley SR:  
XX  
DR WPI: 2001-417975/44.  
DR P-PSDB: AAU03499.  
XX  
PT Modulating angiogenesis in a mammal for treating diseases mediated by  
PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
PT peripheral tissue, by administering antagonist or agonist of TWEAK  
PT receptor  
XX  
PS Example 1; Page 39-40; 46pp; English.  
XX  
CC The sequence represents a DNA from the expression vector  
CC pPC409-12-TWEAK, which encodes a fusion protein comprising a growth  
CC hormone leader, a leucine zipper multimerisation domain, and the  
CC extracellular domain of human TWEAK. The fusion protein was used in  
CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
CC from a COS cell human cDNA library. The TWEAK protein is a  
CC member of the tumour necrosis factor (TNF) family and induces  
CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
CC used in the treatment and diagnosis of human disease. The disorders  
CC mediated by angiogenesis include ocular disorders characterised by ocular  
CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,  
CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
CC rubeosis, uveitis, macular degeneration and corneal graft  
CC neovascularisation, and inflammatory diseases such as arthritis,  
CC rheumatism and psoriasis. Other treatable diseases include malignant and  
CC metastatic conditions such as sarcomas and carcinomas, benign tumours and  
CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
CC scleroderma, vascular adhesions, atherosclerotic plaque  
CC neovascularisation, telangiectasia, wound granulation, coronary  
CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
XX  
SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:  
  
Query Match 42.7%; Score 498.8; DB 22; Length 898;  
Best Local Similarity 87.0%; Pred. No. 3.6e-125;  
Matches 548; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
  
QY 56 AGCTGGGAGACTGGGAGACGCTGTCGCCAGAGCCCTTTCAGAGAGCTGACAGCA 115  
DB 250 AGTTGGGGAGACCGGCGATCGCTCCGCCAGAGCTGCCAGAGAGACTGTGCA 309  
QY 116 GAGGACCGCGGAGCCGCCCTGAGTGAATCCAGACAGAGAGAGAGAGAGATGTGTA 175  
DB 310 GAGGAGACCGAGACCGCTGCGACTGAATCCAGACAGAGAGAGAGAGATGTGCG 369  
QY 176 CCTTCTTGGAACTAGTCCGGCTCGAAGAGTGTCTCTAAAGCCGAGAGCGCGG 235  
DB 370 CCTTCTTGAACCGACTAGTTCGGCTCGCAGAGTGCACCTAAAGCCGAGAGCGG 429  
QY 236 CCTGCGGAGCTATGAGAGCCCATATGAGTGTATCTCGGAGAGAGAGAGAGAGCA 295  
DB 430 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489  
QY 296 CAACGAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 355  
DB 490 CAGCAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549  
QY 356 CCTGCGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 415  
DB 550 CCTTGGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609  
QY 416 CTGACTGTGAGTGCAGTTCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 475  
DB 610 CTGACTGTGAGTGCAGTTCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669  
QY 476 GTGAAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 535  
DB 670 GTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729  
QY 536 CTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 595

DB 730 CTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 789  
QY 596 TCCCTTGGAGATCCGACACCTTCCTGGGCTCATCTTAAGGCTGCCCTTCCTACTAC 655  
DB 790 TCCCTGGGAGATCCGACACCTTCCTGGGCTCATCTTAAGGCTGCCCTTCCTACTAC 849  
QY 656 TTTGAGCTCTTCAAGTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685  
DB 850 TTGAGCTCTTCAAGTTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879  
  
RESULT 12  
AAT22190  
ID AAT22190 standard; cDNA to mRNA; 282 BP.  
XX  
AC AAT22190;  
XX  
DT 27-AUG-1996 (first entry)  
XX  
DE Human gene signature HUMG503761.  
XX  
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KW human; cloning; mapping; non-biased library; diagnosis; detection;  
KW cell typing; abnormal cell function; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9514772-A1.  
XX  
PD 01-JUN-1995.  
XX  
PF 11-NOV-1994; 94WO-JP01916.  
XX  
PR 12-NOV-1993; 93JP-0355504.  
XX  
PA (MATS/) MATSUBARA K.  
PA (OKUB/) OKUBO K.  
PI Matsubara K, Okubo K;  
XX  
DR WPI: 1995-206931/27.  
XX  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissues  
XX  
PS Claim 1; Page 1067; 2245pp; Japanese.  
XX  
CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC given in AAT19001-726837 and which is able to hybridise to part of  
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.  
XX  
SQ Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other:  
  
Query Match 7.1%; Score 82.8; DB 16; Length 282;  
Best Local Similarity 80.6%; Pred. No. 1.9e-12;  
Matches 158; Conservative 0; Mismatches 26; Indels 12; Gaps 5;  
  
QY 969 CAGGCACTGAGATGGGCTGAGACTG-CTGGCAGAGAGAGAGAGAGAGAGAGAGAGCA 1027

||||| ||| ||||||||| | ||||||||| ||| ||| |||||||  
Db 88 CAGGACTAAGAGGGGCTGAGACCTNTGCGGAGGAAGCAAGACTGGGCTAGGACA 147  
Oy 1028 GAAGTTCACACTGTGAGGGGGAAGAGCTGCGGACAGCTCCTCGA----TCCCTG 1083  
Db 148 GAGTTCCCAATNTAGGGGGGAGA-AACAAGACAAGCTCCTCCCTGAGAATTCCTCG 206  
Oy 1084 TGCATTTTGAAA--AGATACATATTTTATTTATTTGTGACAAATGTT---AAATGGA 1137  
Db 207 TCGATTTTAAAAACAGATATTTATTTTNTATTTGTGACAAATGTTGNTAAATGGGA 266  
Oy 1138 TATTAAGAGAAATATA 1153  
Db 267 TATTAATAGAAATAA 282  
RESULT 13  
ABK29540  
ID ABK29540 standard; cDNA; 195 bp.  
AC ABK29540;  
XX 23-APR-2002 (first entry)  
Df  
XX Colon adenocarcinoma-specific cDNA #66.  
DE  
XX Human: colon adenocarcinoma; colon cancer; tumour; gene; ss.  
KW  
XX Homo sapiens.  
OS  
XX WO200196389-A2.  
PN  
XX 20-DEC-2001.  
PD  
XX 07-JUN-2001; 2001WO-US18574.  
PF  
XX 09-JUN-2000; 2000US-210667P.  
PR  
XX 22-NOV-2000; 2000US-252614P.  
XX  
XX (CORI-) CORIXA CORP.  
PA  
XX Meagher MJ, King GE, Xu J, Secrist H;  
PI  
XX WPI: 2002-098052/13.  
DR  
XX  
XX New isolated polynucleotide encoding a polypeptide comprising a portion  
PT of colon tumour protein, for detection, diagnosis and therapy of human  
PT colon cancer  
PS  
XX Claim 1: Page 133; 211pp; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) encoding a  
CC polypeptide (II) comprising a portion of a colon tumour protein. A new  
CC oligonucleotide (III) that hybridises to (I) is useful for  
CC determining the presence of a cancer in a patient. (II) or antigen  
CC presenting cells expressing (I) are useful for stimulating and/or  
CC expanding T cells specific for a tumour protein, by contacting T cells  
CC with (I), (II) or antigen-presenting cells that express (I). (I), (II),  
CC or antigen presenting cells that express (II) are useful for treating  
CC colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated  
CC from a patient with (I), (II), or antigen presenting cells that express  
CC (II), so that T cells proliferate, and administering to the patient an  
CC effective amount of the proliferated T cells, thus inhibiting the  
CC development of a cancer in the patient. A new composition is useful for  
CC stimulating an immune response in a patient. (I) or (II) is useful in  
CC vaccines and pharmaceutical compositions for prevention and treatment of  
CC colon cancer and for the diagnosis and monitoring of the cancers. (I),  
CC (II) or an antibody against (II) is useful for detection, diagnosis and/  
CC or therapy of human colon cancer. (I) is useful as a probe or primer for  
CC nucleic acid hybridisation, and in the design and preparation of ribozyme  
CC molecules for inhibiting expression of (II) in tumour cells. ABK29475-  
CC ABK29851 represent human colon adenocarcinoma-specific cDNA sequences of  
CC the invention.

XX  
SQ Sequence 195 bp; 49 A; 51 C; 58 G; 37 T; 0 other;  
Query Match 6.0%; Score 69.8; DB 24; Length 195;  
Best Local Similarity 83.5%; Pred. No. 5,4e-09;  
Matches 91; Conservative 0; Mismatches 17; Indels 1; Gaps 1;  
Oy 969 CAGGACTAGATGGGCTGTGACCTGTGTGCAGAAAGCCAGAGAACTGTGAGTACGACG 1028  
Db 69 CAGGACTAAGAGGGGCTGTGACCTGTGTGCAGAAAGCCAGAAAGACTGGGCTAGGCCAG 128  
Oy 1029 AAGTTCACACTGTGAGGGGGAAGAGCTGCGGACAGACTCCTCGA 1077  
Db 129 GAGTTCACCAATGTGAGGGGCGAGA-AACAAGACAAGCTCCTCCCTGA 176  
RESULT 14  
ABN55975  
ID ABN55975 standard; DNA; 65 bp.  
XX  
XX ABN55975;  
AC  
XX 15-JUL-2002 (first entry)  
Df  
XX Mouse spliced transcript detection oligonucleotide SEQ ID NO:28723.  
DE  
XX  
XX Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
KW  
XX Mus musculus.  
OS  
XX WO200210449-A2.  
PN  
XX 07-FEB-2002.  
PD  
XX 20-JUL-2001; 2001WO-IB01903.  
PF  
XX 28-JUL-2000; 2000US-221607P.  
PR  
XX 02-MAY-2001; 2001US-287724P.  
XX  
XX (COMP-) COMPUGEN INC.  
PA  
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
PI  
XX WPI: 2002-257383/30.  
DR  
XX  
XX New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes  
PS  
XX Example 1: SEQ ID 28723; 47pp; English.  
XX  
XX The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the  
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises  
CC several oligonucleotides, each capable of hybridising selectively to a  
CC set of messenger RNAs transcribed from a given transcription unit of  
CC the genome, which encodes one or more messenger RNA splice variants.  
CC The oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterising the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a  
CC particular biological or pathological state, and so allowing the  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition; to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN59589 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:58:15 ; Search time 3683.27 Seconds

(without alignments)  
9228.782 Million cell updates/sec

Title: US-09-245-198a-1

Perfect score: 1168

Sequence: 1 ggtcgtgagctcgtgctcgtg.....ataatcatgattctcttc 1168

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb\_da:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rnd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1168	100.0	1168	10	AF030100 Mus muscu
2	711	60.9	203083	2	AC069459 Mus muscu
3	711	60.9	234182	10	AL603707 Mouse DNA
4	628.6	53.8	1353	6	AX201324 Sequence
5	628.6	53.8	1368	6	AF055872 Homo sapi
6	624	53.4	1306	6	AF030099 Homo sapi
7	597.8	51.2	1236	6	AR140407 Sequence
8	501	42.9	138792	2	AC119115 Rattus no
9	498.8	42.7	898	6	AX180714 Sequence
10	409.4	35.1	1651	9	BC019047 Homo sapi
11	304	26.0	215795	2	AC127470 Pan trogl
12	303.8	26.0	177703	2	AC016876 Homo sapi
13	278.2	23.8	153553	2	AC126921 Bos tauru
14	237.4	20.3	177553	2	AC130192 Sus scrof
15	212	18.2	161428	2	AC126925 Canis fam
16	156.8	13.4	184026	2	AC098923 Rattus no
17	111.6	9.6	203281	2	AC126237 Canis fam
18	88.4	7.6	7218	6	166494 Sequence 14
19	86.4	7.4	139405	2	AC126239 Felis cat
20	69.8	6.0	190376	6	AX379024 Sequence
21	61	5.2	100376	2	AC123372 Rattus no
22	58.4	5.0	64424	2	AC124103 Mus muscu
23	56.2	4.8	114260	2	AC098956 Rattus no
24	56.2	4.8	125020	9	AF429315 Homo sapi
25	55.8	4.8	80112	2	AC128800 Rattus no
26	55.4	4.7	69119	2	AC128983 Rattus no
27	55.4	4.7	241370	2	AC094063 Rattus no
28	55	4.7	185822	2	AC073554 Homo sapi
29	54.8	4.7	100511	2	AC010774 Homo sapi
30	54.6	4.7	221758	2	AC068947 Mus muscu
31	54.4	4.7	87120	2	AC012225 Homo sapi
32	54.2	4.6	298166	2	AC087563 Homo sapi
33	54	4.6	111627	2	AC108626 Rattus no
34	54	4.6	184036	2	AC126109 Rattus no
35	54	4.6	310371	2	AC096296 Rattus no
36	53.8	4.6	99517	2	AC131402 Rattus no
37	53.6	4.6	205350	2	AC078946 Mus muscu
38	53.4	4.6	146986	2	AC095995 Rattus no
39	53.4	4.6	203281	2	AC126237 Canis fam
40	53.2	4.6	49430	2	AC100434 Mus muscu
41	53.2	4.6	182269	2	AC107416 Rattus no
42	53	4.5	55061	2	AC091597 Mus muscu
43	53	4.5	176645	2	AC114115 Rattus no
44	53	4.5	187252	10	AL607109 Mouse DNA
45	52.8	4.5	84514	10	AL627264 Mouse DNA

# ALIGNMENTS

RESULT 1

AF030100 1168 bp mRNA 1linear ROD 20-DEC-1997

LOCUS AF030100

DEFINITION Mus musculus TWEAK mRNA, partial cds.

ACCESSION AF030100

VERSION AF030100.1 GI:2707220

KEYWORDS

SOURCE

Mus musculus.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Chicheportliche, Y., Bourdon, P. R., Xu, H., Hsu, Y. M., Scott, H., Hession, C., Garcia, I., and Browning, J. L.

REFERENCE 1 (bases 1 to 1168)

AUTHORS

TWEAK, a new secreted ligand in the tumor necrosis factor family





In the feature table with their source databases: Em., EMBL, Sw., SWISSPROT, Tr., TREMBL, Wp., WORMPEP, information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-422L16 is from the RPI-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

This sequence is the entire insert of clone RP23-422L16.

## FEATURES

source

1..234182  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone\_lib="RP23-422L16"  
/clone\_1lb="RP23-422L16"  
complement(84050..84131)  
/note="Sequence from uni-directional primer reads and dGTP big dye terminator reads only."  
misc\_feature  
BASE COUNT 59310 a 56824 c 57519 g 60529 t  
ORIGIN

Query Match 60.9%; Score 711; DB 10; Length 234182;  
Best Local Similarity 99.6%; Pred. No. 2.6e-176;  
Matches 744; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

OY 425 CAGGTGACCTTTGATGAGGAAGGCTGTCTACCTGAGCTGAGCTTGCTGTAACGCT 484  
|||||  
Db 6623 CAGGTGACCTTTGATGAGGAAGGCTGTCTACCTGAGCTGAGCTTGCTGTAACGCT 66264  
OY 485 GTGCTGGCCCTGGCTGCTGGAAGAAATTCATGACCAACAGCAAGCTTCTGCGGCC 544  
|||||  
Db 66263 GTGCTGGCCCTGGCTGCTGGAAGAAATTCATGACCAACAGCAAGCTTCTGCGGCC 66204  
OY 545 CAGCTCCGTTTGGCCAGAGTGTCTGGCTGTTCGCCCTGGCGGCAAGGCTTCCCTTGG 604  
|||||  
Db 66203 CAGCTCCGTTTGGCCAGAGTGTCTGGCTGTTCGCCCTGGCGGCAAGGCTTCCCTTGG 66144  
OY 605 ATCCGACACCTCCCTGGGCTCATCTTAAAGCTGCCCCCTTCTTAACTACTTGGACTC 664  
|||||  
Db 66143 ATCCGACACCTCCCTGGGCTCATCTTAAAGCTGCCCCCTTCTTAACTACTTGGACTC 66084  
OY 665 TTTTCAAGTTCAGTGAAGGGGCTTGTCTCCAGATTCCTTAACTTCCCTGCTCCAGG 724  
|||||  
Db 66083 TTTTCAAGTTCAGTGAAGGGGCTTGTCTCCAGATTCCTTAACTTCCCTGCTCCAGG 66024  
OY 725 AGATACACACACCTCCCTGACCCACCCACCTCCCTCCCTGCTGCTGCTGCTGCTG 784  
|||||  
Db 66023 AGATACACACACCTCCCTGACCCACCCACCTCCCTCCCTGCTGCTGCTGCTGCTG 65964  
OY 785 CAGTCCCTGT-CTCTCCCAAGGAGCAGCAGAGCTTGTTCAGATG-TTTCATTTCCACAGA 842  
|||||  
Db 65963 CAGTCCCTGTCTCTCTCAAGGAGCAGCAGAGCTTGTTCAGATGTTTTCATTTCCACAGA 65904  
OY 843 CGATCCCTGTCTCTTC-TTAACTCCATCCACACCAACTATTCACCTCAGTCTGCC 901  
|||||  
Db 65903 CGATCCCTGTCTCTTC-TTAACTCCATCCACACCAACTATTCACCTCAGTCTGCC 65844  
OY 902 AAGCCCTCTATATCCCTGACCTCCCAACCCACCTCAACCCGACAGCTGTTTATTAAGT 961  
|||||  
Db 65843 AAGCCCTCTATATCCCTGACCTCCCAACCCACCTCAACCCGACAGCTGTTTATTAAGT 65784  
OY 962 TGTGACACAGCAGTGAAGTGGCTGAGCTGTGTGAGAGCAAGCAAGCTTCCCTGATGCC 1021  
|||||  
Db 65783 TGTGACACAGCAGTGAAGTGGCTGAGCTGTGTGAGAGCAAGCAAGCTTCCCTGATGCC 65724  
OY 1022 AGCCAGAAAGTCCCACTGTGAAGGGGAAAGAGCTGGGGAACAAGCTTCCCTGATGCC 1081  
|||||  
Db 65723 AGCCAGAAAGTCCCACTGTGAAGGGGAAAGAGCTGGGGAACAAGCTTCCCTGATGCC 65664  
OY 1082 TGTGAGATTTGAAAGATATATTTTATATATATATGTGCAAAAGTATAAGGATAT 1141  
|||||  
Db 65663 TGTGAGATTTGAAAGATATATTTTATATATATGTGCAAAAGTATAAGGATAT 65604

OY 1142 AAGAGATTAATCATGATTTCTCTTC 1168  
|||||  
Db 65603 AAGAGATTAATCATGATTTCTCTTC 65577

RESULT 4  
LOCUS AX201324 1353 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 3 from Patent WO0153486.  
ACCESSION AX201324  
VERSION AX201324.1 GI:15391154  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 1353)  
Ashkenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L.,  
Hillan, K.J., Marsters, S.A., Pan, J., Pitt, R.M., Roy, M.A., Smith, V.,  
Stone, D.M., Watanabe, C.K. and Wood, W.I.  
Compositions and methods for the treatment of tumour  
Patent: WO 0153486-A 3 26-JUL-2001;  
Genentech, Inc. (US)

## FEATURES

source

1..1353  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 257 a 443 c 389 g 264 t  
ORIGIN

Query Match 53.8%; Score 628.6; DB 6; Length 1353;  
Best Local Similarity 76.2%; Pred. No. 7.5e-155;  
Matches 933; Conservative 0; Mismatches 219; Indels 73; Gaps 10;

OY 2 GTCTGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 61  
|||||  
Db 130 GCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 189  
OY 62 GGGAGCGGGCAAGCGCTGTCTGCGCAGAGGCTTCTCAGAGAGCTGACAGAGAGAG 121  
|||||  
Db 190 GGGAGCGGGCAAGCGCTGTCTGCGCAGAGGCTTCTCAGAGAGCTGACAGAGAGAG 249  
OY 122 CGCGGAGAGGCGGCTGAGTGAATCCAGACAGAGAGAGAGAGAGATGTGATCTTTC 181  
|||||  
Db 250 GACGAGAGCGGCTGAGTGAATCCAGACAGAGAGAGAGAGAGATGTGATCTTTC 309  
OY 182 TTGGAACAACTAGTCCGGCTCGACAGAGTGTCTTAAAGGCGGAGAGGCGGCTCGC 241  
|||||  
Db 310 CTGAAACCACTAGTGTGGCTCGACAGAGTGTCTTAAAGGCGGAGAGAGGCGGCTCGA 369  
OY 242 CGAGTATTGACAGCCCATTAATGAGTTTCACTCCGCGCAGAGAGAGATGAGAGCA 301  
|||||  
Db 370 AGAGCATGCGAGCCCATTAATGAGTTTCACTCCGCGCAGAGAGATGAGAGCA 429  
OY 302 GGTGTGATGGGACAGTGAAGTGGCTGGGAGAGAGCAAAATCAACAGCTCCAGCCCTGT 361  
|||||  
Db 430 GGTGTGATGGGACAGTGAAGTGGCTGGGAGAGAGCAAAATCAACAGCTCCAGCCCTGT 489  
OY 362 CGCTAGAGCGGCAAGATTTGGGGAATTTACAGTATATAGGGCTGGCTCTACTACTGTAC 421  
|||||  
Db 490 CGCTAGAGCGGCAAGATTTGGGGAATTTAGTACCCGGGCTGGCTCTACTACTGTAC 549  
OY 422 TGTGAGTGCATTTGATGAGGAAAGGCTGTCTACTGTAAGCTGAGCTTGTGAGAC 481  
|||||  
Db 550 TGTGAGTGCATTTGATGAGGAAAGGCTGTCTACTGTAAGCTGAGCTTGTGAGAC 609  
OY 482 GGTGTGCTGGCCCTGGCTGGCTGGAGAAATTCAGACCAACAGCAAGCAAGCTTCTCTGG 541  
|||||  
Db 610 GGTGTGCTGGCCCTGGCTGGCTGGAGAAATTCAGACCAACAGCAAGCAAGCTTCTCTGG 669  
OY 542 CCCAGCTCCGTTTGTGCGAGGTGTGTGGCTGTTCGCGCTGGCGGCAAGGCTTCTCTCC 601  
|||||







QY 838 -----ACAGACGATTCCTTGTCTTCTTAACATCCATCCACCACTATCCACCTC 891  
 DB 925 ACATAAATACAGTATTCACACTCTATCTTAACACTCCGCCACGCCACTCTCCACCTC 984  
 QY 892 ACTAGCTCCCAAGCCCTAC-----TTATCCCTACTCCGCCACCTC 936  
 DB 985 ACTAGCTCCCAAGCCCTAC-----TTATCCCTACTCCGCCACCTC 1044  
 QY 937 CACCGACGAGCTGTTTATGACTTGTGAC----- 968  
 DB 1045 CAGACCCCAAGGAGCTGTTTATGACTTGTGAC----- 1104  
 QY 969 -----CAGGCACTGAGATGGGCTGAGCCGCTGAGAGAGCAAGCAAGCTGGGACTAG 1023  
 DB 1105 CACTTACGAGCACTAGAGAGGGCTGAGCCCTGGGCGCAGAGCAAGCAAGCTGGGCTAG 1164  
 QY 1024 GCCAGAGCTCCCAACTGTGAGGGGGAAGAGCTGGGACAGCTCCCTCCGGA---TC 1079  
 DB 1165 GCCAGAGCTCCCAACTGTGAGGGGGAAGAGCTGGGACAGCTCCCTCCGGA---TC 1223  
 QY 1080 CCGTGAGATTTGAAA--AGATACATTTTATTTATTTATGACAAAATGT---TAAAT 1134  
 DB 1224 CCGTGAGATTTTAAACAGATATTTTATTTATTTATTTATGACAAAATGTGATTAAT 1283  
 QY 1135 GGATATTAAGAGATTAATCA 1156  
 DB 1284 GGATATTAAGATTAATCA 1305

RESULT 7  
 LOCUS ARI140407 1236 bp DNA Linear PAT 16-JUN-2001  
 DEFINITION Sequence 1 from patent US 6207642.  
 ACCESSION ARI140407 GI:14482903  
 VERSION ARI140407.1  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1236)  
 AUTHORS Wiley,S.R.  
 TITLE Member of the TNF family useful for treatment and diagnosis of disease  
 JOURNAL Patent: US 6207642-A 1 27-MAR-2001;  
 FEATURES Location/Qualifiers  
 source 1..1236  
 BASE COUNT 225 a 416 c 358 g 237 t  
 ORIGIN

Query Match 51.2% Score 597.8; DB 6; Length 1236;  
 Best Local Similarity 75.2%; Pred. No. 1e-146;  
 Matches 879; Conservative 0; Mismatches 222; Indels 68; Gaps 8;  
 QY 2 GTGTGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 61  
 DB 73 GGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 132  
 QY 62 GGGAGCTGGGCAAGCTGTCTGCGCAGAGAGCTTCTCAGAGAGAGCTGACAGCAGAGAC 121  
 DB 133 GGGAGCTGGGCAAGCTGTCTGCGCAGAGAGAGCTTCTCAGAGAGAGCTGACAGCAGAG 192  
 QY 122 CGCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 181  
 DB 193 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252  
 QY 182 TTGGAACACTAGTCTGGGCTGGAGAGAGTGTCTCTAAAGCGGAGAGCGGCGCTGCG 241  
 DB 253 CTGAACCGAGTGTGGGCTGGAGAGAGTGTCTCTAAAGCGGAGAGCGGCGCTGCG 312  
 QY 242 CGAGCTATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301  
 DB 313 AGAGGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372

QY 302 GGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361  
 DB 373 GGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432  
 QY 362 CGCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421  
 DB 433 CGCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492  
 QY 422 TGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481  
 DB 493 TGTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552  
 QY 482 GGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541  
 DB 553 GGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612  
 QY 542 CCCAGCTCCGTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601  
 DB 613 CCCAGCTCCGTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672  
 QY 602 CGGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 661  
 DB 673 CGGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732  
 QY 662 CCGTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721  
 DB 733 CCGTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792  
 QY 722 AGGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
 DB 793 -----CGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 847  
 QY 781 GGTCCAGTCTGTCTCTCC--TCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837  
 DB 848 GGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 907  
 QY 838 -----ACAGAGATTCCTGCTCTTAACTATCCATCCACAGAGAGAGAGAGAGAG 891  
 DB 908 ACATTAATACAGATTTTCCACTCTTATCTTAAACAGAGAGAGAGAGAGAGAGAGAG 967  
 QY 892 ACTAGCTCCCAAGCCCTAC-----TTATCCCTACTCCGCCACCTC 936  
 DB 968 ACTAGCTCCCAAGCCCTAC-----TTATCCCTACTCCGCCACCTC 1027  
 QY 937 CACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 968  
 DB 1028 CAGACCCCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1087  
 QY 969 -----CAGGCACTGAGATGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1023  
 DB 1088 CACTTACGAGCACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1147  
 QY 1024 GCCAGAGATTCCAACTGTGAGGGGGAAGAGCTGGGACAGAGCTCCCTCCGGA---TC 1079  
 DB 1148 GCCAGAGATTCCAACTGTGAGGGGGAAGAGCTGGGACAGAGAGAGAGAGAGAGAGAG 1206  
 QY 1080 CCGTGAGATTTGAAAAGATATTTT 1108  
 DB 1207 CCGTGAGATTTTAAACAGATATTTT 1235  
 RESULT 8  
 LOCUS AC119115 138792 bp DNA Linear HTG 18-JUL-2002  
 DEFINITION Rattus norvegicus clone CH230-32ON23. \*\*\* SEQUENCING IN PROGRESS  
 ACCESSION AC119115  
 VERSION AC119115.2 GI:21746718  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 138792)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alstbrooks,S.L., Amaralunga,H.C., Ayele,M., Banks,T.,  
Barbieri,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,  
Bouch,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,S.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabali,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Correll,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kralovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lileu,C., Liu,J., Liu,W., Louiseged,H.,  
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivas,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G.,  
Scherer,S., Scott,G., Shen,H., Shoshtrali,N., Sisson,I.,  
Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
COMMENT

Unpublished  
2 (bases 1 to 138792)  
Moriy K.C.  
Direct Submission  
Submitted (25-APR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 138792)  
Worley K.C.  
Direct Submission  
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20303440.

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
Project Information  
Center project name: GING  
Center clone name: CH230-320N23  
Sequencing vector: Plasmid  
Sequencing method: Big Dye  
Sequencing program: Phrap  
Consensus quality: 112720 bases at least Q40  
Consensus quality: 11666 bases at least Q30  
Consensus quality: 119165 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 32 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1022: contig of 1022 bp in length  
1023 1122: gap of unknown length  
1123 2394: contig of 1272 bp in length  
2395 2494: gap of unknown length  
2495 3663: contig of 1169 bp in length  
3664 4879: gap of unknown length  
4880 4979: gap of unknown length  
4980 6787: contig of 1808 bp in length  
6788 8192: contig of 1305 bp in length  
8193 8292: gap of unknown length  
8293 9461: contig of 1169 bp in length  
9462 9561: gap of unknown length  
9562 11435: contig of 1874 bp in length  
11436 11535: gap of unknown length  
11536 13535: contig of 2000 bp in length  
13536 13635: gap of unknown length  
13636 17233: contig of 3688 bp in length  
17234 17424: gap of unknown length  
17425 20406: contig of 2963 bp in length  
20407 20506: gap of unknown length  
20507 23445: contig of 2939 bp in length  
23446 23545: gap of unknown length  
23546 25600: contig of 2055 bp in length  
25601 25700: gap of unknown length  
25701 28049: contig of 2349 bp in length  
28050 30474: gap of unknown length  
30475 30575: gap of unknown length  
30576 35498: contig of 4924 bp in length  
35499 35598: gap of unknown length  
35599 39284: contig of 3666 bp in length  
39285 39385: gap of unknown length  
39386 41970: contig of 2586 bp in length  
41971 42070: gap of unknown length  
42071 45659: contig of 3589 bp in length  
45660 45759: gap of unknown length  
45760 50332: contig of 4573 bp in length  
50333 50432: gap of unknown length  
50433 54695: contig of 4263 bp in length  
54696 54795: gap of unknown length  
54796 58289: contig of 3494 bp in length  
58290 58390: gap of unknown length  
58391 63599: contig of 5210 bp in length  
63600 63700: gap of unknown length  
63701 70305: contig of 6606 bp in length  
70306 70405: gap of unknown length  
70406 76123: contig of 5718 bp in length  
76124 76223: gap of unknown length  
76224 84961: contig of 8738 bp in length  
84962 85061: gap of unknown length  
85062 93614: contig of 8553 bp in length  
93615 93715: gap of unknown length  
93716 103352: contig of 9638 bp in length  
103353 103452: gap of unknown length  
103453 110299: contig of 6847 bp in length  
110300 110399: gap of unknown length  
110400 116180: contig of 5781 bp in length  
116181 116280: gap of unknown length  
116281 125484: contig of 9204 bp in length  
125485 125584: gap of unknown length



	RESULT 10
LOCUS	BC019047
DEFINITION	Homo sapiens, similar to tumor necrosis factor (ligand) superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA.
ACCESSION	BC019047
VERSION	BC019047.1 GI:17512138
KEYWORDS	MGC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1651)
REFERENCE	Strausberg,R. Direct Submission Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
AUTHORS	NIH-MGC Project URL: http://mgc.nci.nih.gov
JOURNAL	Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Louis Staudt CDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland, Web site: http://www.nisc.nih.gov/ nisc.mgc@nihgtl.nih.gov
REMARK	Contact: Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R., Llan,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McGloskey,J.C., McDowell,J., Pearson,R., Snyder,B., Stancilpop,S., Thomas,P.J., Tongson,E.E., Touchman,J.W., Tsurgason,C., Vogt,J.L., Walker,M.A., Zhang,L.-H. and Green,E.D.
COMMENT	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 30 Row: P Column: 5 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596. Location/Qualifiers 1..1651 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:20669 IMAGE:4766071" /tissue_type="Primary B-Cells from Tonsils" /clone_id="NIH_MGC_48" /lab_host="DH10B-R" /note="Vector: pOTB7" 106..510 /codon_start=1 /product="Similar to tumor necrosis factor (ligand) superfamily, member 12" /protein_id="AAH19047.1" /db_xref="GI:17512139" /translation="MAARSRQRRRGRGPEPTALLVPLAGIGLAICTGLILAVVISI GSRASISAOPEAOEIVAEEDDDSEFLNPDEESODAPFLLNRLLVRPRRSAPKPKRTTRTARRIAAHYEHHPRGDGDGAADGGTYTCMP"
FEATURES	
source	
CDS	
BASE COUNT	344 a 517 c 481 g 309 t
ORIGIN	
Query Match	35.1%; Score 409.4; DB 9; Length 1651;
Best Local Similarity	73.7%; Pred. No. 6.le-97;
Matches 685; Conservative	0; Mismatches 171; Indels 73; Gaps 10;
298 AGCAGGTGTGATGGACAGTGATGGCTGGAGAAGACAACAAATCAACAGCTCCAGCCC	357

Db	713	ACCAAGTGTGGACGGAGCAGTAGAGTGGCTGGGAGGAAAGCCAGAAATTCACAGCTCCAGCCC	772
Qy	358	TCCTGCGCTACGACCGCCGACGATTTGGGGAATTTCACATCATCAGGGCTGGGCTCTACTACT	417
Db	773	TCCTGGGCTTACCAACCGCCGAGATCGGGGAGTTATATATACACCGGGCTGGGCTCTACTACT	832
Qy	418	GTACGCTAGAGTGCACCTTGTGATGAGAGGAAAGGCTCTACCTGAAAGCTGAGACTTCTGGT	477
Db	833	GTACGTAGAGTGCACCTTGTGATGAGAGGAAAGGCTCTACCTGAAAGCTGAGACTTCTGGT	892
Qy	478	GAAAGGTGTGCTGGCCCTTGCCCTGCGCTGCGTGGAGAAATTCACAGCACAGCAGACTCTCC	537
Db	893	GGATGTGTGCTGGCCCTTGCCCTGCGCTGCGTGGAGAAATTCACAGCACAGCTGCGGAGTTCCCT	952
Qy	538	TGGGGCCGACGCTCGGTTGTGCGCAGAGTCTGAGGCTGTGGCGGTGGGGGCAAGGGTCTTC	597
Db	953	CGGGCCCGACGCTCGGCTCTGCGCAGAGTCTGAGGCTGTGGCGGTGGGGGCAAGGGTCTTC	1012
Qy	598	CCCTTCGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCGCCCTCTTAACCTACTT	657
Db	1013	CCCTGGGATCCGACACCTCCCTGGGCTCATCTTAAGGCTGCGCCCTCTTAACCTACTT	1072
Qy	658	TGAGCTCTTTCAAGTTCACTGAGGGGCTTGTCTCTCCAGATTCCTTAACTTTCCCTGG	717
Db	1073	CGGACTCTTCCAGGTTCACTGAGGGGCTTGTCTCTCCAGATTCCTTAACTTTCCCTGG	1132
Qy	718	CTCCAGAGCATCACACACCTCCCTTACCCACCCGACCTCTCCAGCCCTCC- GCTGCT	776
Db	1133	CTCC- -----CCTCCAGACACTCTGTGGGACCCCGTCCCTCTGGCCCAACCTCAGCCGCT	1187
Qy	777	CTTGTGTCAGTCTCTGTCTTCC- TCAAAGGACCCAGAGCTGTTCACATGTTTCAT	834
Db	1188	CTTGTGTCAGAGCTCGCCCTCTCTAGAGGCTGCCCTGTTCACAGTGTTCATGTTTTC	1247
Qy	835	TCC-----ACAGACGTATCCTTGTCTCTTAAATCCATCCATCCACACACATATCA	887
Db	1248	TCCACATAAATACAGTATTTCCACTCTTATCTTACACTCCCCACCGCCACTCTCCA	1307
Qy	888	CCTCAGTACGTCCCAAGACCCCTTAC-----TTATCCGTACTCCGCCACG	932
Db	1308	CCTCAGTACGTCCCAATCCCTGACCTTTGAGGCCCCCAGAGATCTGACCTCCCCCTG	1367
Qy	933	CACATCACCCGACACAGTATTATGACTTGTGTGAC-----	968
Db	1368	GCCACAGACCCCGAGGAGTGTGTCTACTGTACTCTGTGGGACAGATGGGTCCAGAG	1427
Qy	969	-----CAGGCACTGAGATGGGCTGAGCTGGTGGCAGAGACCCAGAGAACCTGGGA	1019
Db	1428	ACCCACCTTACAGGCACTAAGAGGGCTGTGACTGGCGGCAAGAAAGCAGATGGGCG	1487
Qy	1020	CTAGGCCAGAAATTTCCCACTGTGAGGGGGAAGAGCTGGGACAAAGCTCTCCCTGGGA--	1077
Db	1488	CTAGGCCAGGAGTTTCCCAATGTGAGGGGCGAGAAACAA-GACAAAGCTCTCCCTGGGA	1546
Qy	1078	--TCCCTGTGATTTTGGAAA--AGATACATATTTTATTTATTTGTGACAAATGT--T	1130
Db	1547	ATTCCCTGTGATTTTAAACACAGATATTTATTTATTTATTTATTTGTGACAAATGTGAT	1606
Qy	1131	AAATGGATATTTAAAGAGAAATTAATCATGA	1159
Db	1607	AAATGGATATTTAAATAGAAATTAATCATGA	1635
RESULT	11		
AC127470			
LOCUS	AC127470		
DEFINITION	Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15		
ACCESSION	AC127470.1		
VERSION	GI:21886866		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	chimpanzee.		
ORGANISM	Pan troglodytes		

REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan. 1 (bases 1 to 215795)  
Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamins, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Latic, P., Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B., Marquies, E.H., Masello, C., Maskeri, B., Mastrian, S.D., McInnes, J.C., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Schueler, M.G., Stantir, J.W., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, D.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 215795)  
Green, E.D.  
Direct Submission  
Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA  
Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc\_zoo@nhgri.nih.gov  
Project Information  
Center project name: cms  
Center clone name: 145D13

Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 203929 bases at least Q40  
Consensus quality: 206865 bases at least Q30  
Consensus quality: 209181 bases at least Q20  
Insert size: 19000; agarose-fp  
Insert size: 214395; sum-of-ctrls  
Quality coverage: 7.30x in Q20 bases; agarose-fp  
Quality coverage: 6.47x in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 3593: contig of 3593 bp in length  
3594 3693: gap of unknown length  
3694 8001: contig of 4308 bp in length  
8002 8101: gap of unknown length  
8102 15639: contig of 7538 bp in length  
15640 15739: gap of unknown length  
15740 24775: contig of 9036 bp in length  
24776 24875: gap of unknown length  
24876 30150: contig of 5275 bp in length  
30151 30251: gap of unknown length  
30251 39212: contig of 8962 bp in length  
39212 39312: gap of unknown length  
39312 49342: contig of 10030 bp in length  
49342 49442: gap of unknown length  
49442 49443: contig of 11514 bp in length  
49443 60956: gap of unknown length  
60956 61056: gap of unknown length  
61056 74520: contig of 13464 bp in length  
74520 74620: gap of unknown length  
74620 89522: contig of 14902 bp in length  
89522 89622: gap of unknown length  
89622 105537: contig of 15915 bp in length  
105537 105637: gap of unknown length  
105637 121197: contig of 15560 bp in length  
121197 121297: gap of unknown length  
121297 138883: contig of 17586 bp in length

FEATURES  
source  
1..215795  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-145D13"  
/clone="11b-RP43"  
1..3593  
/note="assembly-fragment"  
3694..8001  
/note="assembly-fragment"  
8102..15639  
/note="assembly-fragment"  
15740..24775  
/note="assembly-fragment"  
24876..30150  
/note="assembly-fragment"  
30251..39212  
/note="assembly-fragment"  
39313..49342  
/note="assembly-fragment"  
49443..60956  
/note="assembly-fragment"  
61057..74520  
/note="assembly-fragment"  
74621..89522  
/note="assembly-fragment"  
89623..105537  
/note="assembly-fragment"  
105638..121197  
/note="assembly-fragment"  
121298..138883  
/note="assembly-fragment"  
138984..166802  
/note="assembly-fragment"  
166903..215795  
/note="assembly-fragment"

BASE COUNT 52160 a 54859 c 53752 g 55513 t 1511 others  
ORIGIN  
Query Match 26.0%; Score 304; DB 2; Length 215795;  
Best Local Similarity 71.7%; Pred. No. 7.9e-69;  
Matches 501; Conservative 0; Mismatches 155; Indels 74; Gaps 11;  
QY 425 CAGTGCACCTTGATGAGGAAAGGCTGTACCTGAGCTGAGTGAACGT 484  
|||||  
DB 199456 CAGTGCACCTTGATGAGGAAAGGCTGTACCTGAGCTGAGTGAACGT 199515  
QY 485 GTGCTGCGCCCTGCGCTGCGTGAAGAATTCACGACAGACAGACAGCTCTGGGCC 544  
|||||  
DB 199516 GTGCTGCGCCCTGCGCTGCGTGAAGAATTCACGACAGACAGCTCTGGGCC 199575  
QY 545 CAGCTCGTTTGGCGAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 604  
|||||  
DB 199576 CAGCTCGTTTGGCGAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 199635  
QY 605 ATCCGACCCCTCCCTGCGCTATCTTAAGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 664  
|||||  
DB 199636 ATCCGACCCCTCCCTGCGGCTATCTTAAGGCTGCGGCTGCGGCTGCGGCTGCGG 199695  
QY 665 TTTCAGTCACTGAGGAGCTGCTGCTCCAGATTCCTTAACCTTCCCTGGCTCAGG 724  
|||||  
DB 199696 TTTCAGTCACTGAGGAGCTGCTGCTCCAGATTCCTTAACCTTCCCTGGCTCAGG 199752  
QY 725 AGATGACCAACCTCCACACCCACCCACCTCCACACCCACCTCCACACCTCCACCT 783  
|||||  
DB 199753 --CTGACAGCTCTTGAGGACCGGCTCTCTGCGCCACCCCTGAGCTGCTTGGCT 199810  
QY 784 CCAGTCTGCTCTCC--TCAAAGGACGACAGGCTTGTTCACATG--TTTCATTCACACA 840  
|||||





```

Db 37234 GTGCTGCCCCCTGCGCTGCGCTGGAGCAATTCACAGCCACTGCGCGGAGTTCCCTCGGGCCC 37195
Oy 545 CAGCTCCGTTTGTGTCAGAGTGTGTCGGGCTTGCCGCTGCGGCGAGGGCTTCCTTCGG 604
Db 37194 CAGCTCCGCTTGTGTCAGAGTGTGTCGGGCTTGCCGCTGCGGCGAGGGCTTCCTTCGG 37135
Oy 605 ATCCGACCCCTCCCTGCGGCTTCATCTTAAGGCTGCCCCCTTCCTTAACCTACTTGTGACTC 664
Db 37134 ATCCGACCCCTCCCTGCGGCTTCATCTTAAGGCTGCCCCCTTCCTTAACCTACTTGTGACTC 37075
Oy 665 TTTCAAGTTCACAGAGGGGCTGCTCTCCAGATTCCTTAACCTTCCTGCTGCTCAG 724
Db 37074 TTCCAGGTTCACAGAGGGGCTGCTCTCCAGATTCCTTAACCTTCCTGCTGCTCAG 37018
Oy 725 AGCATCACCACACACCTCCATCCCGCCAGCTCCCTCCAGCTCCAGCTCCAGCTCCAGCT 783
Db 37017 --CCTCGACAGCTCTCTGGGCGCGGCTCCCTGCGGCTCCAGCTCCAGCTCCAGCT 36960
Oy 784 CCAGTCTCTGCTCTCTCC--TCAAGGCGACGACAGCTTCCTTCATGTTTCCATTC-- 837
Db 36959 CCAGACCTGCCCCCTCCCTCTGAGGCTGCTGGGCTGCTTCACGTTTTCATCCACACA 36900
Oy 838 ---ACAGACGATCTCTCTCTCTTAAACATCCCATCCACACACTATCCACTACT 894
Db 36899 TAAATACAGTATTCCTCTATCTTACAACTCCCGACCGCCACCTCTCCACTCACT 36840
Oy 895 AGCTCCCAAGGCCCTAC-----TTATCCCTGACCTCCCGACCCACTCAG 939
Db 36839 AGCTCCCAAGGCCCTGAGCCCTTTGAGGCCCCAGTGAATCTGACCTCCCGCTGCGCACAG 36780
Oy 940 CCGACACGCTGTTTATGACTTTTGTGCAC----- 968
Db 36779 ACCCCAGGAGGATTTGTTCACCTGACCTGTGGGCAAGGATGGTCCAGAGACCCAC 36720
Oy 969 ---CAGGACTGATGCGCTGAGCTGTGTCAGAGACGACGACGCTGGAGTGGCC 1026
Db 36719 TTCAGGCACTAAGAGGGCTGACCTGGCGGAGGACCAAGAGACTGCGCTAGGCG 36660
Oy 1027 AGAAGTTCACACCTGAGGGGAGAGAGCTGGGACAAAGCTCTCCCTGGA----TCCCT 1082
Db 36659 AGGAGTTCACCAATGTGAGGGGCGAGA--AACAGACAAAGCTCTCCCTGAGAAATTCCT 36601
Oy 1083 GTGATTTTGA---AGATACTATTTTATTTATTTATTTGACAAATGT---TAAATGA 1137
Db 36600 GTGATTTTGA---AGATACTATTTTATTTATTTATTTATTTGACAAATGTGTAATGA 36541
Oy 1138 TATTAAGAGAAATTAATCATGATTTCTCT 1166
Db 36540 TATTAAGAGAAATTAATCATGATTTCTCT 36512

RESULT 13
AC126921 153553 bp DNA linear HTG 10-JUL-2002
LOCUS AC126921
DEFINITION Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
ACCESSION AC126921.1 GI:21724098
VERSION AC126921
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 153553)
AUTHORS Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-T., Idol, J.R., Karlins, E., Latic, P.,
Lee-Jin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Maguiles, E.H., Mastiello, C., Maskeri, B., Mastrian, S.D.,
McCluskey, J.C., McDowell, J., Paguirigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantirip, S., Thomas, J.W.,

```

```

TITLE Thomas, P. J., Touchman, J. W., Tsurgoun, C., Vogt, J. L., Walker, M. A.,
JOURNAL Wehrby, K. D., Wiggins, L., Young, A., Zhang, L. H. and Green, E. D.
REFERENCE NISC Comparative Sequencing Initiative
AUTHORS 2 (bases 1 to 153553)
TITLE Unpublished
JOURNAL 2 (bases 1 to 153553)
COMMENT Direct Submission
Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
-----
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehgr@nih.gov
-----
Project Information
Center project name: ddi
Center clone name: 045D24
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 14606 bases at least Q40
Consensus quality: 147748 bases at least Q30
Consensus quality: 148824 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 152453; sum-of-ctnigs
Quality coverage: 8.80x in Q20 bases; agarose-fp
Quality coverage: 8.72x in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2312: contig of 2312 bp in length
2313 2412: gap of unknown length
2413 5841: contig of 3429 bp in length
5842 5941: gap of unknown length
5942 8435: contig of 2494 bp in length
8436 8535: gap of unknown length
8536 15799: contig of 7264 bp in length
15799 15899: gap of unknown length
15899 25224: contig of 9325 bp in length
25224 25225: gap of unknown length
25225 32504: contig of 7180 bp in length
32504 32604: gap of unknown length
32604 40970: contig of 8366 bp in length
40970 40971: gap of unknown length
40971 41071: contig of 15520 bp in length
41071 56591: gap of unknown length
56591 73769: contig of 17079 bp in length
73769 73869: gap of unknown length
73869 90859: contig of 16990 bp in length
90859 90860: gap of unknown length
90860 111428: contig of 20469 bp in length
111428 111528: gap of unknown length
111528 11529: contig of 42025 bp in length.
Location/Qualifiers
1. 153553
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-45D24"
/clone_id="RP42"
1. 2312
/note="assembly-fragment"
2413. 5841
/note="assembly-fragment"
clone_end=77
vector_side="left"
5942. 8435
misc_feature

```





Consensus quality: 153125 bases at least Q20  
Insert size: 152000; agarose-fp  
Insert size: 160028; sum-of-ctrls  
Quality coverage: 5.36x in Q20 bases; agarose-fp  
Quality coverage: 5.10x in Q20 bases; sum-of-ctrls

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 4734: contig of 4734 bp in length  
\* 4735 4834: gap of unknown length  
\* 4835 7325: contig of 2491 bp in length  
\* 7326 7425: gap of unknown length  
\* 7426 1261: contig of 4836 bp in length  
\* 1262 1261: gap of unknown length  
\* 1262 19031: contig of 6670 bp in length  
\* 19032 19031: gap of unknown length  
\* 19132 27849: contig of 8718 bp in length  
\* 27850 27949: gap of unknown length  
\* 31801 31901: contig of 3852 bp in length  
\* 31902 31901: gap of unknown length  
\* 40782 40881: contig of 8880 bp in length  
\* 40882 49783: contig of 8902 bp in length  
\* 49784 49883: gap of unknown length  
\* 49884 59150: contig of 9267 bp in length  
\* 59151 59250: gap of unknown length  
\* 59251 68679: contig of 9429 bp in length  
\* 68680 68779: gap of unknown length  
\* 68780 78984: contig of 10205 bp in length  
\* 78985 79085 79084: gap of unknown length  
\* 93447 93446: contig of 14362 bp in length  
\* 93447 93546: gap of unknown length  
\* 102176 102175: contig of 8629 bp in length  
\* 102276 131685: contig of 29410 bp in length  
\* 131686 131785: gap of unknown length  
\* 131786 161428: contig of 29643 bp in length.

## FEATURES

source  
1. 161428  
/organism="Canis familiaris"  
/db\_xref="taxon:9615"  
/clone\_11b="RP81"  
1. 4734  
/note="assembly\_fragment"  
misc\_feature  
4835..7325  
/note="assembly\_fragment"  
misc\_feature  
7426..12261  
/note="assembly\_fragment"  
misc\_feature  
12362..19031  
/note="assembly\_fragment"  
misc\_feature  
19132..27849  
/note="assembly\_fragment"  
misc\_feature  
27950..31801  
/note="assembly\_fragment"  
misc\_feature  
31902..40781  
/note="assembly\_fragment"  
misc\_feature  
40882..49783  
/note="assembly\_fragment"  
misc\_feature  
49884..59150  
/note="assembly\_fragment"  
misc\_feature  
59251..68679  
/note="assembly\_fragment"  
misc\_feature  
68780..78984  
/note="assembly\_fragment"  
misc\_feature  
79085..93446  
/note="assembly\_fragment"

misc\_feature 93547..102175  
/note="assembly\_fragment"  
misc\_feature 102276..131685  
/note="assembly\_fragment"  
misc\_feature 131786..161428  
/note="assembly\_fragment"  
BASE COUNT 39586 a 41863 c 41233 g 37340 t 1406 others  
ORIGIN

Query Match 18.2%; Score 212; DB 2; Length 161428;  
Best Local Similarity 60.9%; Pred. No. 1.5e+44;  
Matches 458; Conservative 0; Mismatches 270; Indels 24; Gaps 6;

Oy 425 CAGGTGCACTTTGATGAGGAAAGGCTGTACTGAGCTGAGCTGCTGTAACGCT 484  
|||||  
Db 42487 CAGGTGCACTTTGATGAGGAAAGGCTGTACTGAGCTGAGCTGCTGTAACGCT 42428  
Oy 485 GTGCTGCGCTGCGCTGCGCTGGAAGAAATCTCAGCCACAGCAGCAGCTCTCCGCGCC 544  
|||||  
Db 42427 GCGCTGCGCTGCGCTGCGCTGGAAGAAATCTCAGCCACAGCAGCAGCTCTCCGCGCC 42368  
Oy 545 CAGCTCGCTTGTGCGCGCTGCTGCGCTGCTGCGCTGCGCTGCGCTGCGCTGCG 604  
|||||  
Db 42367 CAGCTCGCTTGTGCGCGCTGCTGCGCTGCTGCGCTGCGCTGCGCTGCGCTGCG 42308  
Oy 605 ATCCGACCCCTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664  
|||||  
Db 42307 ATCCGACCCCTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42248  
Oy 665 TTTCAGTTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724  
|||||  
Db 42247 TTTCAGTTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42188  
Oy 725 AGCATC-ACCAACATCTCCATCCCAACCCCACTCCCTGCTGCTGCTGCTGCTGCT 783  
-|||  
Db 42187 CTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42128  
Oy 784 CCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843  
|||||  
Db 42127 ATATCTCTTCTTCTTCTTCTGACACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42068  
Oy 844 GTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903  
|||||  
Db 42067 CTGTTGAGGCGCCCATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42015  
Oy 904 AGCCCTACTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963  
|||||  
Db 42014 CACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 41955  
Oy 964 TGCACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
|||||  
Db 41954 GTCCCTTAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 41895  
Oy 1021 --TAGCCAGAAAGTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
|||||  
Db 41894 ACGGAGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 41842  
Oy 1078 TCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134  
|||||  
Db 41841 CCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 41782  
Oy 1135 GGATATTAAGAGAAATATCATGATTTCTCT 1166  
|||||  
Db 41781 GGATATTAAGAGAAATATTAACACAGTCTCTCT 41750

Search completed: March 31, 2003, 11:08:11  
Job time : 6031.27 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 11:50:36 ; Search time 66 Seconds  
(without alignments)  
573.381 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARRPLPLPSLQ.....PMAHLKAAPFLYFGLEQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 65 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	100.0	284	19 AAM47525	Homo sapiens tumou
2	249	87.7	249	20 AAM47525	Human tumour necro
3	249	87.7	249	21 AAM47525	Amino acid sequenc
4	249	87.7	249	21 AAM47525	Human PRO207 antic
5	249	87.7	249	23 AAM47525	Human PRO207 poly
6	241	84.9	249	19 AAM47525	TNF related endoth
7	241	84.9	249	19 AAM47525	Human TREPA (TNF r
8	207	72.9	273	22 AAM47525	TWEAK extracellular
9	146	51.4	146	22 AAM47525	Human TREPA (TNF r
10	143	50.4	189	19 AAM47525	TNF related endoth

11	143	50.4	189	22 AAM47525	Human UL4flap TRP
12	107	37.7	208	20 AAM47525	Human TNRL3 protei
13	46	16.2	211	20 AAM47525	Mouse TNRL3 protei
14	32	11.3	225	19 AAM47525	Mus musculus tumou
15	32	11.3	225	19 AAM47525	Amino acid sequenc
16	9	3.2	58	21 AAM47525	Human secreted pro
17	9	3.2	58	21 AAM47525	Human secreted pro
18	9	3.2	748	20 AAM47525	Extended amino aci
19	9	3.2	749	23 AAM47525	M vaccae GVS-9 pro
20	20	2.8	54	22 AAM47525	Protonibacterium
21	8	2.8	55	22 AAM47525	Peptide #10143 enc
22	8	2.8	55	22 AAM47525	Protein #7982 enco
23	8	2.8	55	22 AAM47525	Human brain expres
24	8	2.8	55	22 AAM47525	Human bone marrow
25	8	2.8	55	22 AAM47525	Peptide #7292 enco
26	8	2.8	55	22 AAM47525	Peptide #10488 enc
27	8	2.8	55	22 AAM47525	Human peptide enco
28	8	2.8	65	22 AAM47525	Peptide #11487 enc
29	8	2.8	65	22 AAM47525	Protein #8889 enco
30	8	2.8	65	22 AAM47525	Human brain expres
31	8	2.8	65	22 AAM47525	Human bone marrow
32	8	2.8	65	22 AAM47525	Peptide #8055 enco
33	8	2.8	65	22 AAM47525	Peptide #11960 enc
34	8	2.8	65	22 AAM47525	Human peptide enco
35	8	2.8	69	22 AAM47525	Novel human diagn
36	8	2.8	71	22 AAM47525	Novel human diagn
37	8	2.8	84	22 AAM47525	Proionibacterium
38	8	2.8	110	22 AAM47525	Novel human diagn
39	8	2.8	117	22 AAM47525	Human polypeptide
40	8	2.8	184	22 AAM47525	Drosophila melanog
41	8	2.8	190	23 AAM47525	Streptococcus poly
42	8	2.8	198	22 AAM47525	Human EST encoded
43	8	2.8	222	23 AAM47525	Human ovarian anti
44	8	2.8	266	22 AAM47525	Novel human diagn
45	8	2.8	307	22 AAM47525	Novel human diagn
46	8	2.8	370	22 AAM47525	Novel human diagn
47	8	2.8	370	22 AAM47525	Novel human diagn
48	8	2.8	370	22 AAM47525	Novel human diagn
49	8	2.8	424	22 AAM47525	Novel human diagn
50	8	2.8	430	22 AAM47525	Human EST encoded
51	8	2.8	454	22 AAM47525	Drosophila melanog
52	8	2.8	586	22 AAM47525	Novel human diagn
53	8	2.8	592	22 AAM47525	Novel human secret
54	8	2.8	603	22 AAM47525	Novel human diagn
55	8	2.8	633	22 AAM47525	Novel human diagn
56	8	2.8	799	22 AAM47525	Novel human diagn
57	7	2.5	12	16 AAM47525	Cell adhesion prot
58	7	2.5	17	23 AAM47525	Human lymphotoxin
59	7	2.5	18	23 AAM47525	Human tumour suppr
60	7	2.5	23	23 AAM47525	Human tumour suppr
61	7	2.5	24	20 AAM47525	Human pre-prohepar
62	7	2.5	25	21 AAM47525	Antigenic peptide
63	7	2.5	44	20 AAM47525	Polypeptide fragme
64	7	2.5	44	20 AAM47525	Human secreted pro
65	7	2.5	44	22 AAM47525	Human gene 14 enco

## ALIGNMENTS

RESULT 1  
AAM47525  
ID AAM47525 standard; Protein: 284 AA.  
XX AAM47525:  
AC AAM47525:  
AC AAM47525:  
DT 21-JUL-1998 (first entry)  
XX  
XX  
DE Homo sapiens tumour necrosis factor related ligand (TRELL).  
KW TRELL: tumour necrosis factor related ligand; tnfr; treatment;  
KW cancer; autoimmune disease; immune system; stimulation; suppression;  
KW graft rejection.

XX Homo sapiens.  
OS  
XX MO9805783-A1.  
PN  
XX 12-FEB-1998.  
PD  
XX 07-AUG-1997; 97MO-US13945.  
PF  
XX 18-MAR-1997; 97US-0040820.  
PR 07-AUG-1996; 96US-0023541.  
PR 18-OCT-1996; 96US-0028515.  
XX  
PA (BIOJ ) BIOGEN INC.  
XX (UYGE-) UNIV GENEVA FACULTY MEDICINE.  
XX  
PI Browning JL, Chicheportliche Y;  
XX  
XX WPI; 1998-145619/13.  
DR N-PSDB; AAV18600.  
DR  
XX  
XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts.  
XX  
XX Claim 12; Pages 50-51; 69pp; English.  
PS  
XX  
XX The sequence is that of human tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon- gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. It's coding sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of gene encoding probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.  
XX  
XX  
SQ Sequence 284 AA;  
Query Match 100.0%; Score 284; DB 19; Length 284;  
Best Local Similarity 100.0%; Pred. No. 2.2e-252;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2  
AAV09369  
ID AAV09369 standard; Protein; 249 AA.  
AC  
XX AAV09369;  
XX  
DT 15-JUL-1999 (first entry)  
XX  
XX Human tumour necrosis factor Apo-3 ligand protein sequence.  
DE  
XX  
XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;  
KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
KW cancer.  
XX  
XX Homo sapiens.  
OS  
XX MO9919490-A1.  
PN  
XX  
XX 22-APR-1999.  
PD  
XX  
XX 09-OCT-1998; 98MO-US21407.  
PF  
XX  
XX 17-DEC-1997; 97US-0069862.  
PR 10-OCT-1997; 97US-0062037.  
PR  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX  
XX Ashkenazi AJ, Marsters SA, Pittl R;  
PI  
XX  
XX WPI; 1999-287982/24.  
DR  
XX  
XX N-PSDB; AAX56000.  
DR  
XX  
XX New human Apo-3 ligand (a tumour necrosis factor) homologue  
PT  
XX  
XX Claim 1; Fig 1; 74pp; English.  
PS  
XX  
XX The present sequence represents a human tumour necrosis factor (TNF)  
CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in  
CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and  
CC to induce JNK/SAPK-dependent responses in mammalian cells.  
CC  
XX  
SQ Sequence 249 AA;  
Query Match 87.7%; Score 249; DB 20; Length 249;  
Best Local Similarity 100.0%; Pred. No. 2.7e-220;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC AAB07526;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Amino acid sequence of a soluble recombinant human TWEAK protein.  
XX  
KW TWEAK protein; immunological disorder; immune response; inflammation;  
KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
KW Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour  
XX  
OS Homo sapiens.  
XX  
PN MO200042073-A1.  
XX  
PD 20-JUL-2000.  
XX  
PE 14-JAN-2000; 2000MO-US01044.  
XX  
PR 15-JAN-1999; 99US-0116168.  
XX  
PA (BIOJ ) BIOGEN INC.  
XX  
PI Rennett P;  
XX  
DR WPI: 2000-476036/41.  
XX  
PT Preventing and treating immune responses using modulators, especially  
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
PT treating e.g. inflammation and graft versus host disease -  
XX  
PS Disclosure; Fig 1; 45pp; English.  
XX  
XX The present sequence represents a TWEAK protein. The specification  
CC describes a method for preventing or treating an immunological  
CC disorder and/or inhibiting an immune response in an animal. The  
CC method comprises administering a TWEAK blocking agent. The method may  
CC be used for preventing and treating immune disorders associated with  
CC inappropriate expression and/or activity of TWEAK. These disorders  
CC include autoimmune diseases, acute and chronic inflammation, organ  
CC transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell  
CC malignancies, septic and other forms of shock, loss of immune  
CC responsiveness (as seen in human immunodeficiency virus (HIV)  
CC infections) and failure of the immune response to tumour growth.  
XX  
XX Sequence 249 AA;

Query Match	87.7%	Score 249	DB 21	Length 249
Best Local Similarity	100.0%	Pred. No. 276-220		
Matches 249	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	36	MAARRSQRRRGRCRGPGTALLVPLALGGLALACGLLAAVSLGSRALSAOEPAOEEL	95	
Db	1	MAARRSQRRRGRCRGPGTALLVPLALGGLALACGLLAAVSLGSRALSAOEPAOEEL	60	
QY	96	VAEEDODSELPNPOTEEODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGOD	155	
Db	61	VAEEDODSELPNPOTEEODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGOD	120	
QY	156	GAQAGVDGTSGVSMPEARINSSSPLEKYNROIIGEFITYTRAGLYLYLCOVHDEGKAYULKD	215	
Db	121	GAQAGVDGTSGVSMPEARINSSSPLEKYNROIIGEFITYTRAGLYLYLCOVHDEGKAYULKD	180	
QY	216	LLVDGVLLRLCRLEEPSATAAASLGGQLRLCOVSGLLALRPGSSLIRTLRPMANLKAAPFL	275	
Db	181	LLVDGVLLRLCRLEEPSATAAASLGGQLRLCOVSGLLALRPGSSLIRTLRPMANLKAAPFL	240	
QY	276	TYFGLEQVH 284		
Db	241	TYFGLEQVH 249		

ID	AA95338	standard; Protein; 249 AA.
XX		
AC	AA95338;	
XX		
DT	25-SEP-2000	(first entry)
XX		
DE	Human PRO207 antitumour protein.	
XX		
KW	PRO207; human; antitumour; tumour; therapy; cytostatic;	
KW	breast cancer; ovarian cancer; renal cancer; colorectal cancer;	
KW	uterine cancer; prostate cancer; lung cancer; bladder cancer;	
KW	central nervous system cancer; melanoma; leukaemia; neoplasm.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FH	Peptide	1..40
FT		/label= Signal_peptide
FT	Protein	41..249
FT		/label= PRO207
FT	Modified-site	27..33
FT		/note= "N-myristoylation"
FT	Modified-site	29..35
FT		/note= "N-myristoylation"
FT	Modified-site	36..42
FT		/note= "N-myristoylation"
FT	Modified-site	45..51
FT		/note= "N-myristoylation"
FT	Modified-site	118..124
FT		/note= "N-myristoylation"
FT	Modified-site	121..127
FT		/note= "N-myristoylation"
FT	Modified-site	125..131
FT		/note= "N-myristoylation"
FT	Modified-site	128..134
FT		/note= "N-myristoylation"
FT	Modified-site	139..143
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	10..14
FT		/note= "amidation"
FT	Modified-site	97..101
FT		/note= "amidation"
FT	Peptide	24..35
FT		/note= "prokaryotic membrane lipoprotein lipid"
XX		
PN	WO200037638-A2.	
PD	29-JUN-2000.	
XX		
PF	02-DEC-1999;	99WO-US28565.
XX		
PR	02-DEC-1998;	98US-0113286.
PR	08-MAR-1999;	99WO-US05022.
PR	21-APR-1999;	99US-0130232.
PR	28-APR-1999;	99US-0131445.
PR	14-MAY-1999;	99US-0134287.
PR	20-JUL-1999;	99US-0144758.
PR	26-JUL-1999;	99US-0145698.
PR	15-SEP-1999;	99WO-US21090.
PR	15-SEP-1999;	99WO-US21547.
XX		
PA	(GETH ) GENENTECH INC.	
XX		
PI	Ashtenazi AJ, Goddard A, Godowski PJ, Gurney AL, Maisters SA,	
PI	Napier KM, Pitti RM, Wood WI;	
XX		
XX	WPI; 2000-442668/38.	
DR	N-PSDB; AAA49717.	
XX		
PT	Novel composition to inhibit neoplastic cell growth or for treating	
PT	tumor in mammal comprises polypeptides PRO179, PRO207, PRO319,	
PT	PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or	
PT	PRO866	

XX Claim 19; Fig 4; 172pp; English.

PS The present sequence is that of human antitumour protein PRO207,

XX as deduced from a foetal kidney cDNA clone (see AAA49117). PRO207

CC shows amino acid sequence identity to tumour necrosis factor family

CC members, especially human lymphotoxin-beta (23.4%) and human CD40

CC ligand (19.8%). Mol.wt. is 27,216. A claimed method for inhibiting

CC the growth of a tumour cell comprises exposing the tumor cell

CC to PRO179, PRO207, PRO320, PRO221, PRO224, PRO328, PRO301,

CC PRO326, PRO356, PRO509 or PRO866 (see AA95337-49), their

CC agonists or chimeric polypeptides incorporating them. The tumour

CC is especially a cancer selected from breast, ovarian, renal,

CC colorectal, uterine, prostate, lung, bladder and central nervous

CC system cancer, melanoma and leukaemia. Methods for the recombinant

CC expression of the antitumour proteins are also provided.

XX

SQ Sequence 249 AA:

Query Match 87.7%; Score 249; DB 21; Length 249;

Best Local Similarity 100.0%; Pred. No. 2.7e-220;

Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARRSQRRRGRGEPGTALLVPLALGCLALACGLLAVVSLGSRASLSAQEPAQEEL 95

DB 1 MAARRSQRRRGRGEPGTALLVPLALGCLALACGLLAVVSLGSRASLSAQEPAQEEL 60

QY 96 VAEEDDPSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 155

DB 61 VAEEDDPSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 120

QY 156 GAAGVDTGTVSGWEARINSSSPLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 215

DB 121 GAAGVDTGTVSGWEARINSSSPLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 180

QY 216 LVDGVLALRCLEFSATASSLGPOLRLCOVSGLLALRPSSLRITLPMHNLKAAPFL 275

DB 181 LVDGVLALRCLEFSATASSLGPOLRLCOVSGLLALRPSSLRITLPMHNLKAAPFL 240

QY 276 TYFGLEFQVH 284

DB 241 TYFGLEFQVH 249

RESULT 5

AAU86129

ID AAU86129 standard; Protein; 249 AA.

XX AAU86129;

XX 15-JUL-2002 (first entry)

XX Human PRO207 polypeptide.

XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;

KW leukaemia; neuronal disorder; stromal disorder; blastococelic disorder;

KW inflammatory disorder; immune disorder; anglogenetic disorder;

KW cytoslastic; neuroprotective.

XX

OS Homo sapiens.

XX

PN WO200153486-A1.

XX

PD 26-JUL-2001.

XX

PF 11-FEB-2000; 2000WO-US03565.

XX

PR 08-MAR-1999; 99WO-US05028.

PR 11-MAR-1999; 99US-123972P.

PR 11-MAY-1999; 99US-133459P.

PR 02-JUN-1999; 99WO-US12252.

PR 22-JUN-1999; 99US-140650P.

PR 22-JUN-1999; 99US-140653P.

PR 20-JUL-1999; 99US-144758P.

PR 26-JUL-1999; 99US-145698P.

PR 28-JUL-1999; 99US-146222P.

PR 17-AUG-1999; 99US-149395P.

PR 31-AUG-1999; 99US-151689P.

PR 01-SEP-1999; 99WO-US20111.

PR 15-SEP-1999; 99WO-US21090.

PR 30-NOV-1999; 99WO-US28313.

PR 01-DEC-1999; 99WO-US28301.

PR 01-DEC-1999; 99WO-US28634.

PR 05-JAN-2000; 2000WO-US00219.

XX

PA (GETH ) GENENTECH INC.

XX

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;

PI Martens SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM,

PI Watanabe CK, Wood WI;

XX

DR WPI, 2002-205567/26.

DR N-PSDB; ABR40255.

XX

PT Thirty five nucleic acids encoding PRO polypeptides, useful for

PT treating benign or malignant tumours, leukaemias and lymphoid

PT malignancies, inflammatory, anglogenetic and immunologic disorders -

PS Claim 61; Fig 4; 302pp; English.

XX

XX The present invention relates to the isolation of novel human PRO

CC polypeptides and the polynucleotide sequences encoding them. The

CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are

CC useful for treating benign or malignant tumours (e.g. renal, kidney,

CC bladder, breast, etc), leukaemias and lymphoid malignancies, other

CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,

CC macropagal, stromal and blastococelic disorders, inflammatory, immune

CC and anglogenetic disorders. The polynucleotide sequences are also

CC useful in gene therapy. AU86128-AU86162 represent the human PRO

CC polypeptides of the invention.

XX

SQ Sequence 249 AA:

Query Match 87.7%; Score 249; DB 23; Length 249;

Best Local Similarity 100.0%; Pred. No. 2.7e-220;

Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARRSQRRRGRGEPGTALLVPLALGCLALACGLLAVVSLGSRASLSAQEPAQEEL 95

DB 1 MAARRSQRRRGRGEPGTALLVPLALGCLALACGLLAVVSLGSRASLSAQEPAQEEL 60

QY 96 VAEEDDPSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 155

DB 61 VAEEDDPSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 120

QY 156 GAAGVDTGTVSGWEARINSSSPLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 215

DB 121 GAAGVDTGTVSGWEARINSSSPLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 180

QY 216 LVDGVLALRCLEFSATASSLGPOLRLCOVSGLLALRPSSLRITLPMHNLKAAPFL 275

DB 181 LVDGVLALRCLEFSATASSLGPOLRLCOVSGLLALRPSSLRITLPMHNLKAAPFL 240

QY 276 TYFGLEFQVH 284

DB 241 TYFGLEFQVH 249

RESULT 6

AAW29745

ID AAW29745 standard; Protein; 249 AA.

XX AAW29745;

XX 27-OCT-1998 (first entry)



DE TNF related endothelium proliferative agent protein.  
 XX  
 KW TNF: endothelium proliferative agent; TREPA: wound healing; cancer;  
 KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 XX  
 OS Homo sapiens.  
 XX MO9835061-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 12-FEB-1998; 98WO-US02859.  
 XX  
 PR 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Willey SR;  
 XX  
 DR WPI: 1998-447255/38.  
 DR N-PSDB: AAV47613.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 XX treatment of autoimmune disease, tumours and inflammation  
 XX  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 XX The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 XX  
 SO Sequence 249 AA;  
 Query Match 84.9%; Score 241; DB 19; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-213;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID AAE00891 standard; Protein; 249 AA.  
 XX  
 AC AAE00891;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 XX  
 KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vulneryary.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 98..249  
 FT /label= Extracellular\_domain  
 XX  
 PN US6207642-B1.  
 XX  
 PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Willey SR;  
 XX  
 DR WPI: 2001-280760/29.  
 DR N-PSDB: AAD04350.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 PS Claim 1; Column 75-76; 53pp; English.  
 XX  
 XX The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is clone ID #690050 human TREPA.  
 CC  
 XX  
 SO Sequence 249 AA;  
 Query Match 84.9%; Score 241; DB 22; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-213;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```

OY 202 VHFDEGKAVYLLKLDLVGVALRCLEEFSSATAASSLGQPLRLCOVSGLLALRPGSSLR 261
DB 107 VHFDEGKAVYLLKLDLVGVALRCLEEFSSATAASSLGQPLRLCOVSGLLALRPGSSLR 166
OY 262 RFLPWAHLKAAPFLTYFGLFOVH 284
DB 167 RFLPWAHLKAAPFLTYFGLFOVH 189

RESULT 12
AAW93590
ID AAW93590 standard: Protein; 208 AA.
AC AAW93590;
XX
XX 18-JUN-1999 (first entry)
DE Human TNRL3 protein.
XX
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KM developmental abnormality; gestational abnormality; prostate cancer;
KM APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KM apoptosis; human.
XX
XX Homo sapiens.
OS
XX
XX W09911791-A2.
PN
XX 11-MAR-1999.
PD
XX 04-SEP-1998; 98WO-US18393.
PF
XX 05-SEP-1997; 97US-0924634.
PR
XX (UNITM ) UNITV WASHINGTON.
PA
XX Chaudhary PM;
PI
XX
XX WPI: 1999-205191/17.
DR N-PSDB; AAX23424.
XX
XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
XX Claim 40; Fig 13A; 156pp; English.
PS
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the changer in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
XX Sequence 208 AA:
XX
XX Query Match 37.7%; Score 107; DB 20; Length 208;

```

```

Best Local Similarity 99.5%; Pred. No. 5.8e-90;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 77 VSLGSRASLSAEPPOEELVAEEDDPSELNPQTESODPAFLNRLVPRRSAPGRKT 136
DB 1 VSLGSRASLSAEPPOEELVAEEDDPSELNPQTESODPAFLNRLVPRRSAPGRKT 60
OY 137 RARRAIAHYEVHPRGODGAGVGTGSGWEARINSSPLRYNRQIGEIFVTRAGLY 196
DB 61 RARRAIAHYEVHPRGODGAGVGTGSGWEARINSSPLRYNRQIGEIFVTRAGLY 120
OY 197 YLYCOVHFDEGKAVYLLKLDLVGVALRCLEEFSSATAASSLGQPLRLCOVSGLLALRPG 256
DB 121 YLYCOVHFDEGKAVYLLKLDLVGVALRCLEEFSSATAASSLGQPLRLCOVSGLLALRPG 180
OY 257 SSLRIRTLPWAHLKAAPFLTYFGLFOVH 284
DB 181 SSLRIRTLPWAHLKAAPFLTYFGLFOVH 208

RESULT 13
AAW93591
ID AAW93591 standard: Protein; 211 AA.
AC AAW93591;
XX
XX 18-JUN-1999 (first entry)
DE Mouse TNRL3 protein.
XX
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KM developmental abnormality; gestational abnormality; prostate cancer;
KM APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KM apoptosis; mouse.
XX
XX Mus sp.
OS
XX
XX W09911791-A2.
PN
XX 11-MAR-1999.
PD
XX 04-SEP-1998; 98WO-US18393.
PF
XX 05-SEP-1997; 97US-0924634.
PR
XX (UNITM ) UNITV WASHINGTON.
PA
XX Chaudhary PM;
PI
XX
XX WPI: 1999-205191/17.
DR N-PSDB; AAX23425.
XX
XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
XX Claim 40; Fig 13B; 156pp; English.
PS
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the changer in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active

```

CC fragments and APO4 signal transducer molecules that specifically interact  
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
CC activity. The method is performed in vivo or in vitro. APO polypeptides  
CC are all useful as immunogens for preparing antibodies. APO4 is also  
CC useful for diagnosis/treatment of developmental or gestational  
CC abnormalities. APO4 was transfected to human breast carcinoma cell line  
CC MCF-7, and induced apoptosis.

XX  
SQ Sequence 211 AA;

Query Match 16.2%; Score 46; DB 20; Length 211;  
Best Local Similarity 100.0%; Pred. No. 6.1e-34;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 239 GPOLRLGVSGSLALRPSSLRITLPMAHLKAPFLTYGLPQVH 284  
Db 166 GPOLRLGVSGSLALRPSSLRITLPMAHLKAPFLTYGLPQVH 211

RESULT 14  
AAW47524

ID AAW47524 standard; Protein: 225 AA.

AC AAW47524;

DT 21-JUL-1998 (first entry)

DE Mus musculus tumour necrosis factor related ligand (TRELL).

XX  
KW TRELL; tumour necrosis factor related ligand; tnf; treatment;  
KW cancer; autoimmune disease; immune system; stimulation; suppression;  
KW graft rejection.

XX  
OS Mus musculus.

XX  
FH Key Location/Qualifiers

FT Domain 1..21 /note="hydrophobic, transmembrane domain"

XX  
PN WO9805783-A1.

XX  
PD 12-FEB-1998.

XX  
PF 07-AUG-1997; 97WO-US13945.

XX  
PR 18-MAR-1997; 97US-0040820.

XX  
PR 07-AUG-1996; 96US-0023541.

XX  
PR 18-OCT-1996; 96US-0028515.

XX  
PA (BIOJ ) BIOGEN INC.

XX  
PI (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX  
PI Browning JL, Chicheportiche Y;

XX  
DR WPI; 1998-145619/13.

XX  
DR N-PSDB; AAV18599.

XX  
PS Claim 12; Pages 48-50; 69pp; English.

XX  
CC The sequence is that of mouse tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon- gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its

CC receptor. It's coding sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of probe probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.

XX  
SQ Sequence 225 AA;

Query Match 11.3%; Score 32; DB 19; Length 225;  
Best Local Similarity 100.0%; Pred. No. 4.7e-21;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 139 RRAIAHYEVHPRGCGAAGVDTYSGWEE 170  
Db 80 RRAIAHYEVHPRGCGAAGVDTYSGWEE 111

RESULT 15  
AAB07527

ID AAB07527 standard; Protein: 225 AA.

AC AAB07527;

DT 20-OCT-2000 (first entry)

DE Amino acid sequence of a soluble recombinant murine TWEAK protein.

XX  
KW TWEAK protein; immunological disorder; immune response; inflammation;  
KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.

XX  
OS Mus sp.

XX  
PN WO200042073-A1.

XX  
PD 20-JUL-2000.

XX  
PF 14-JAN-2000; 2000WO-US01044.

XX  
PR 15-JAN-1999; 99US-0116168.

XX  
PA (BIOJ ) BIOGEN INC.

XX  
PI Renner P;

XX  
DR WPI; 2000-476036/41.

XX  
PT Preventing and treating immune responses using modulators, especially  
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
PT treating e.g. inflammation and graft versus host disease -  
XX  
PS Disclosure; Fig 1; 45pp; English.

XX  
CC The present sequence represents a TWEAK protein. The specification  
CC describes a method for preventing or treating an immunological  
CC disorder and/or inhibiting an immune response in an animal. The  
CC method comprises administering a TWEAK blocking agent. The method may  
CC be used for preventing and treating immune disorders associated with  
CC inappropriate expression and/or activity of TWEAK. These disorders  
CC include autoimmune diseases, acute and chronic inflammation, organ  
CC transplant rejection, graft-versus-host disease (GVHD), lymphoid cell  
CC malignancies, septic and other forms of shock, loss of immune  
CC responsiveness (as seen in human immunodeficiency virus (HIV)  
CC infections) and failure of the immune response to tumour growth.

XX  
SQ Sequence 225 AA;

Query Match 11.3%; Score 32; DB 21; Length 225;  
Best Local Similarity 100.0%; Pred. No. 4.7e-21;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 RRAIAHVEVHPRPGDGAQAGVDGTVSGMEE 170  
 DB 80 RRAIAHVEVHPRPGDGAQAGVDGTVSGMEE 111

RESULT 16  
 ID AAG01265 standard; Protein; 58 AA.  
 AC AAG01265;  
 DT 06-OCT-2000 (first entry)  
 DE Human secreted protein, SEQ ID NO: 5346.  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KM gene therapy; chromosome mapping.  
 OS Homo sapiens.  
 PN EPI033401-A2.  
 PD 06-SEP-2000.  
 XX 21-FEB-2000; 2000EP-0200610.  
 PF 26-FEB-1999; 99US-0122487.  
 PR (GEST ) GENSET.  
 PA Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 DR N-PSDB; AAC01271.  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX Claim 13; SEQ ID 5346; 71pp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

SQ Sequence 58 AA:  
 Query Match 3.2%; Score 9; DB 21; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSLLDPEIS 9  
 DB 46 MSLLDPEIS 54

RESULT 17  
 ID AAG01266 standard; Protein; 58 AA.  
 AC AAG01266;  
 XX

DT 06-OCT-2000 (first entry)  
 DE Human secreted protein, SEQ ID NO: 5347.  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KM gene therapy; chromosome mapping.  
 OS Homo sapiens.  
 PN EPI033401-A2.  
 PD 06-SEP-2000.  
 XX 21-FEB-2000; 2000EP-0200610.  
 PF 26-FEB-1999; 99US-0122487.  
 PR (GEST ) GENSET.  
 PA Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 DR N-PSDB; AAC01272.  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX Claim 13; SEQ ID 5347; 71pp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

SQ Sequence 58 AA:  
 Query Match 3.2%; Score 9; DB 21; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSLLDPEIS 9  
 DB 46 MSLLDPEIS 54

RESULT 18  
 ID AAY14906 standard; protein; 748 AA.  
 AC AAY14906;  
 DT 25-OCT-1999 (first entry)  
 DE Extended amino acid sequence for GVS-9.  
 XX Mycobacterium vaccae protein; antigen; T cell activation; cytokine;  
 KM dendritic cell maturation; infectious disease; immune disorder; cancer;  
 KM respiratory system; mycobacterial infection; allergy; tuberculosis;  
 KM leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;  
 KM dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;  
 KM squamous cell carcinoma; melanoma.  
 XX Mycobacterium vaccae.

XX MO9932634-A2.  
 PN 01-JUL-1999.  
 PD 23-DEC-1998; 98WO-N200189.  
 PF 04-DEC-1998; 98US-0205426.  
 PR 23-DEC-1997; 97US-0996624.  
 PR 23-DEC-1997; 97US-0997080.  
 PR 23-DEC-1997; 97US-0997362.  
 PR 11-JUN-1998; 98US-0095855.  
 PR 17-SEP-1998; 98US-0156181.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;  
 PI WPI; 1999-430163/36.  
 DR N-PSDB; AA211368.  
 DR  
 XX  
 XX Enhancing immune response to an antigen  
 PS Claim 1; Page 209-210; 243pp; English.  
 XX  
 XX The invention provides heat-killed Mycobacterium vaccae, or recombinant  
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate  
 CC T cells and natural killer cells, to stimulate the production of  
 CC cytokines, to enhance the expression of co-stimulatory molecules on  
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation  
 CC and function. The proteins can be expressed by standard recombinant  
 CC methodology. Pharmaceutical compositions comprising the proteins or  
 CC nucleic acid sequences encoding the proteins can be used for the  
 CC treatment, prevention, and detection of disorders including infectious  
 CC diseases, immune disorders and cancer. In particular, the compounds and  
 CC methods are used for treatment of diseases of the respiratory system,  
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,  
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as  
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,  
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell  
 CC carcinoma and melanoma.  
 CC  
 XX  
 SQ Sequence 748 AA;  
 Query Match 3.2%; Score 9; DB 20; Length 748;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 60 ALGIGLALA 68  
 ||||||||  
 DB 282 ALGIGLALA 290  
 RESULT 19  
 ABB73512  
 ID ABB73512 standard; Protein; 749 AA.  
 XX  
 AC ABB73512;  
 XX  
 XX 08-APR-2002 (first entry)  
 XX  
 DE M vaccae GVs-9 protein SEQ ID NO: 154.  
 XX  
 DE  
 XX  
 KW Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;  
 KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;  
 KW antipsoriatic; dermatological; antiinflammatory; antiallergic;  
 KW Th2 immune response; immunomodulatory.  
 XX  
 XX  
 OS Mycobacterium vaccae.  
 XX  
 PI US6328978-B1.  
 PN  
 XX  
 PD 11-DEC-2001.

XX 02-JUN-1999; 99US-0324542.  
 PF 23-DEC-1997; 97US-0997080.  
 PR (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 PA Watson JD, Tan PLJ, Prestidge R;  
 PI WPI; 2002-138361/18.  
 DR N-PSDB; ABL36274.  
 XX  
 XX Inhibiting skin inflammation associated with skin disorder e.g.  
 PT psoriasis, by administering composition comprising delipidated and  
 PT deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae  
 PT culture filtrate -  
 XX  
 PS Example 6; Column 161-164; 116pp; English.  
 XX  
 XX The present invention relates to a method of inhibiting skin inflammation  
 CC associated with a skin disorder selected from psoriasis, atopic  
 CC dermatitis and allergic contact dermatitis, which involves administering  
 CC a composition containing delipidated and deglycolipidated Mycobacterium  
 CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be  
 CC treated may also include alopecia areata, and skin cancers such as basal  
 CC cell carcinoma, squamous cell carcinoma and melanoma. The composition  
 CC acts by inhibiting the Th2 immune response. The present sequence is a  
 CC protein described in the exemplification of the invention.  
 CC  
 XX  
 SQ Sequence 749 AA;  
 Query Match 3.2%; Score 9; DB 23; Length 749;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 60 ALGIGLALA 68  
 ||||||||  
 DB 282 ALGIGLALA 290  
 RESULT 20  
 AAU51863  
 ID AAU51863 standard; Protein; 54 AA.  
 XX  
 AC AAU51863;  
 XX  
 DT 27-FEB-2002 (first entry)  
 XX  
 DE Propionibacterium acnes immunogenic protein #12759.  
 XX  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 PN  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US12865.  
 XX  
 XX 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 PN  
 XX  
 DR WPI; 2001-616774/71.

DR N-PSDB; AAS59552.  
XX  
PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -  
XX  
PS Example 1; SEQ ID NO 13058; 1069pp; English.  
XX  
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 54 AA:  
  
Query Match 2.8%; Score 8; DB 22; Length 54;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 11 RLPLPRS 18  
DB 21 RLPLPRS 28  
  
RESULT 21  
ABBA42637  
ID ABBA42637 standard; Peptide; 55 AA.  
XX  
AC ABBA42637;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #10143 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.

XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
XX  
PS Claim 27; SEQ ID NO 35272; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 55 AA:  
  
Query Match 2.8%; Score 8; DB 22; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 61 LGLGLALA 68  
DB 12 LGLGLALA 19  
  
RESULT 22  
ABB25983  
ID ABB25983 standard; Protein; 55 AA.  
XX  
AC ABB25983;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #7982 encoded by probe for measuring heart cell gene expression.  
XX  
KW Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200157274-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00666.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488899/53.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
PS Claim 15; SEQ ID NO 27753; 530pp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.



CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 55 AA;

Query Match 2.8%; Score 8; DB 22; Length 55;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGLGLALA 68  
Db 12 LGLGLALA 19

RESULT 23

AAM63528  
ID AAM63528 standard; Protein; 55 AA.

AC AAM63528;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35633.

KM Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
XX epilepsy; cancer.

OS Homo sapiens.

PN W0200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

PS Example 4; SEQ ID NO: 35633; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system,  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is a protein encoded by one of  
CC the probes of the invention.

XX  
SQ Sequence 55 AA;

Query Match 2.8%; Score 8; DB 22; Length 55;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGLGLALA 68  
Db 12 LGLGLALA 19

RESULT 24

AAM76342  
ID AAM76342 standard; Protein; 55 AA.

AC AAM76342;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36648.

KM Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma.

OS Homo sapiens.

PN W0200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 36648; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.

XX  
SQ Sequence 55 AA;

OY 61 LGLGLALA 68

Db 12 LGLGLALA 19

RESULT 25

AAM20858  
ID AAM20858 standard; Protein; 55 AA.

AC AAM20858;

DT 12-OCT-2001 (first entry)

DE Peptide #7292 encoded by probe for measuring cervical gene expression.

```

XX  Probe: human; microarray; gene expression; cervical epithelial cell;
KW  cervical cancer.
XX
XX  Homo sapiens.
XX
XX  WO200157278-A2.
XX
XX
XX  09-AUG-2001.
XX
XX  30-JAN-2001; 2001WO-US00670.
XX
XX  04-FEB-2000; 2000US-0180312.
XX  26-MAY-2000; 2000US-0207456.
XX  30-JUN-2000; 2000US-0608408.
XX  03-AUG-2000; 2000US-0632366.
XX  21-SEP-2000; 2000US-0234687.
XX  27-SEP-2000; 2000US-0236359.
XX  04-OCT-2000; 2000GB-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX  Penn SG, Hanzel DK, Chen W, Rank DR;
XX  WPI; 2001-488901/53.
XX
XX  Human genome-derived single exon nucleic acid probes useful for
XX  analyzing gene expression in human cervical epithelial cells -
XX
XX  Claim 27; SEQ ID No 25684; 487bp; English.
XX
XX  The present invention relates to human single exon nucleic acid probes
XX  (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded
XX  by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX  can be used to produce a single exon microarray, which can be used for
XX  measuring human gene expression in a sample derived from human cervical
XX  epithelial cells. By measuring gene expression, the probes are therefore
XX  useful in grading and/or staging of diseases of the cervix, notably
XX  cervical cancer.
XX  Note: The sequence data for this patent did not form part of the printed
XX  specification, but was obtained in electronic format directly from WIPO
XX  at ftp.wipo.int/pub/published_pct_sequences.
XX
XX  Sequence 55 AA:
SQ
Query Match 2.8%; Score 8; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 61 LGIGLALA 68
DB 12 LGIGLALA 19

```

```

PF 30-JAN-2001; 2001WO-US00663.
XX
XX  04-FEB-2000; 2000US-0180312.
XX  26-MAY-2000; 2000US-0207456.
XX  30-JUN-2000; 2000US-0608408.
XX  03-AUG-2000; 2000US-0632366.
XX  21-SEP-2000; 2000US-0234687.
XX  27-SEP-2000; 2000US-0236359.
XX  04-OCT-2000; 2000GB-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX  Penn SG, Hanzel DK, Chen W, Rank DR;
XX  WPI; 2001-488897/53.
XX
XX  Human genome-derived single exon nucleic acid probes useful for
XX  analyzing gene expression in human placenta -
XX
XX  Claim 27; SEQ ID No 36720; 654bp; English.
XX
XX  The present invention relates to single exon nucleic acid probes (SENP:
XX  see AAI31315-AA157546). The present sequence is a peptide encoded by one
XX  such probe. The probes are useful for producing a microarray for
XX  predicting, measuring and displaying gene expression in samples derived
XX  from human placenta. The probes are useful for antenatal diagnosis of
XX  human genetic disorders.
XX
XX  Sequence 55 AA:
SQ
Query Match 2.8%; Score 8; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 61 LGIGLALA 68
DB 12 LGIGLALA 19

```

```

RESULT 27
ABG45635
ID ABG45635 standard; Peptide; 55 AA.
XX
XX  ABG45635;
XX
XX  19-AUG-2002 (first entry)
XX
XX  Human peptide encoded by genome-derived single exon probe SEQ ID 35300.
XX
XX  Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX  chronic obstructive pulmonary disease; interstitial lung disease;
XX  familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX  tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX  Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX  pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
XX  pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX  primary ciliary dyskinesia; pulmonary hypertension;
XX  hyaline membrane disease.
XX
XX  Homo sapiens.
XX
XX  WO200186003-A2.
XX
XX  15-NOV-2001.
XX
XX  30-JAN-2001; 2001WO-US00665.
XX
XX  04-FEB-2000; 2000US-180312P.
XX  26-MAY-2000; 2000US-207456P.
XX  30-JUN-2000; 2000US-0608408.
XX  03-AUG-2000; 2000US-0632366.
XX  21-SEP-2000; 2000US-234687P.
XX  27-SEP-2000; 2000US-234687P.
XX  27-SEP-2000; 2000US-236359P.

```

PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2002-114183/15.  
 XX  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 XX  
 PS Claim 27: SEQ ID NO 35300; 634pp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridise at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberculous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic,  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 55 AA;  
 XX  
 QY  
 Db 61 LGGLGALA 68  
 12 LGGLGALA 19  
 11111111  
 Query Match 2.8%; Score 8; DB 23; Length 55;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Peptide #11487 encoded by human foetal liver single exon probe.  
 XX  
 KM Human: foetal liver: gene expression; single exon nucleic acid probe.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632386.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 XX  
 PS Claim 27: SEQ ID NO 36616; 639pp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 65 AA;  
 XX  
 QY  
 Db 61 LGGLGALA 68  
 22 LGGLGALA 29  
 11111111  
 Query Match 2.8%; Score 8; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 29  
 ABB26890  
 ID ABB26890 standard; Protein: 65 AA.  
 XX  
 AC ABB26890;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Protein #8889 encoded by probe for measuring heart cell gene expression.  
 XX  
 KW Human: gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00666.

```
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID NO 28660; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease.
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 65;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 LGTGLALA 68
|||
Db 22 LGTGLALA 29
XX
RESULT 30
AAM64984
ID AAM64984 standard; Protein: 65 AA.
XX
AC AAM64984;
XX
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 37089.
XX
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
```

```
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 37089; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 65 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 65;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 LGTGLALA 68
|||
Db 22 LGTGLALA 29
XX
RESULT 31
AAM77706
ID AAM77706 standard; Protein: 65 AA.
XX
AC AAM77706;
XX
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38012.
XX
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 38012; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
```

CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.

XX  
SQ Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
| | | | | | | |  
Db 22 LGIGLALA 29

RESULT 32  
AAM21621  
ID AAM21621 standard; Protein; 65 AA.

XX  
AC AAM21621;

XX  
DT 12-OCT-2001 (first entry)

XX  
DE Peptide #8055 encoded by probe for measuring cervical gene expression.

XX  
KM Probe: human; microarray; gene expression; cervical epithelial cell;  
KM cervical cancer.

XX  
OS Homo sapiens.

XX  
PN W0200157278-A2.

XX  
PD 09-AUG-2001.

XX  
PE 30-JAN-2001; 2001MO-US00670.

XX  
PR 04-FEB-2000; 2000US-0180312.

XX  
PR 26-MAY-2000; 2000US-0207456.

XX  
PR 30-JUN-2000; 2000US-0608408.

XX  
PR 03-AUG-2000; 2000US-0632366.

XX  
PR 21-SEP-2000; 2000US-0234687.

XX  
PR 27-SEP-2000; 2000US-0236359.

XX  
PR 04-OCT-2000; 2000GB-0024263.

XX  
PS (MOLE-) MOLECULAR DYNAMICS INC.

XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
DR WPI; 2001-488901/53.

XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -

XX  
PS Claim 27; SEQ ID No 26447; 487pp; English.

XX  
CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs: see A110068-A128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
| | | | | | | |  
Db 22 LGIGLALA 29

RESULT 33  
AAM37923  
ID AAM37923 standard; Protein; 65 AA.

XX  
AC AAM37923;

XX  
DT 17-OCT-2001 (first entry)

XX  
DE Peptide #11960 encoded by probe for measuring placental gene expression.

XX  
KM Probe: microarray; human; placenta; antenatal diagnosis;  
KM genetic disorder.

XX  
OS Homo sapiens.

XX  
PN W0200157272-A2.

XX  
PD 09-AUG-2001.

XX  
PE 30-JAN-2001; 2001MO-US00663.

XX  
PR 04-FEB-2000; 2000US-0180312.

XX  
PR 26-MAY-2000; 2000US-0207456.

XX  
PR 30-JUN-2000; 2000US-0608408.

XX  
PR 03-AUG-2000; 2000US-0632366.

XX  
PR 21-SEP-2000; 2000US-0234687.

XX  
PR 27-SEP-2000; 2000US-0236359.

XX  
PR 04-OCT-2000; 2000GB-0024263.

XX  
PS (MOLE-) MOLECULAR DYNAMICS INC.

XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
DR WPI; 2001-48897/53.

XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -

XX  
PS Claim 27; SEQ ID No 38192; 654pp; English.

XX  
CC The present invention relates to single exon nucleic acid probes (SENPs:  
CC see A113135-AA157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.

XX  
SQ Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
| | | | | | | |  
Db 22 LGIGLALA 29

RESULT 34  
ABG46737  
ID ABG46737 standard; Peptide; 65 AA.

XX  
AC ABG46737;

XX  
DT 19-AUG-2002 (first entry)

XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 36402.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW Chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tubercous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemostiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX Homo sapiens.  
 OS WO200186003-A2.  
 FN 15-NOV-2001.  
 XX 30-JAN-2001; 2001WO-US000665.  
 XX 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SC, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2002-114183/15.  
 XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 PS Claim 27; SEQ ID No 36402; 634pp; English.  
 XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridise at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tubercous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
 CC haemostiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
 CC and hyaline membrane disease. The present sequence is a peptide/protein

CC encoded by a single exon probe of the invention.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp://ipo.int/pub/published\_pct\_sequences.  
 CC XX  
 SO Sequence 65 AA:  
 Query Match 2.8%; Score 8; DB 23; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 61 LGLGLALA 68  
 |||||  
 22 LGLGLALA 29  
 RESULT 35  
 ABG03723  
 ID ABG03723 standard; Protein; 69 AA.  
 XX  
 AC ABG03723;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #3714.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX WO200175067-A2.  
 XX 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US08631.  
 PF 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS67910.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 PS Claim 20; SEQ ID No 34082; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG0377 represent novel human  
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 69 AA;

Query Match 2.8%; Score 8; DB 22; Length 69;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
| | | | | | | |  
Db 16 LGIGLALA 23

RESULT 36

ABG03663  
ID ABG03663 standard; Protein; 71 AA.

AC ABG03663;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #3654.

KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.  
N-PSDB; AAS67850.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

PS Claim 20; SEQ ID No 34022; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 71 AA;

Query Match 2.8%; Score 8; DB 22; Length 71;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
| | | | | | | |  
Db 16 LGIGLALA 23

RESULT 37

AAU61282  
ID AAU61282 standard; Protein; 84 AA.

AC AAU61282;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #22178.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.  
N-PSDB; AAS59615.

PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -  
XX Example 1; SEQ ID No 22477; 106pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 84 AA;  
Query Match 2.8%; Score 8; DB 22; Length 84;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 259 LRIRTLPM 266  
| | | | | | | |  
DB 34 LRIRTLPM 41  
RESULT 38  
ABG20259  
ID ABG20259 standard; Protein; 110 AA.  
XX  
AC ABG20259;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #20250.  
XX  
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KM food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI: 2001-639362/73.  
DR N-PSDB; AAS84446.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 20; SEQ ID NO 50618; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridization probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 110 AA;  
Query Match 2.8%; Score 8; DB 22; Length 110;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 61 LGLGLALA 68  
| | | | | | | |  
DB 44 LGLGLALA 51  
RESULT 39  
AAO08094  
ID AAO08094 standard; Protein; 117 AA.  
XX  
AC AAO08094;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 21986.  
XX  
KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KM tissue growth factor; immunomodulatory; cancer; leukaemia;  
KM nervous system disorders; arthritis; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI: 2001-514838/56.  
DR N-PSDB; AAI88025.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 20; SEQ ID NO 21986; 1399bp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 117 AA;  
Query Match 2.8%; Score 8; DB 22; Length 117;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 61 LGLGLALA 68  
| | | | | | | |



Db 54 LGIGIALA 61

## RESULT 40

ABB67712

ID ABB67712 standard; Protein: 184 AA.

XX ABB67712;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 29928.

KM Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN MO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001MO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR N-PSDB; ABL11815.

DR WPI: 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 29928; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 184 AA;

Query Match 2.8%; Score 8; DB 22; Length 184;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 LLLAVVSL 79

|||||

Db 16 LLLAVVSL 23

Search completed: March 31, 2003, 14:13:36  
Job time : 71 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:07:46 ; Search time 39 Seconds  
(Without alignments)  
700.056 Million cell updates/sec

Title: US-09-245-198A-4  
Perfect score: 284  
Sequence: 1 MSLLDFEISARRLPLPRSLG.....PMAHLKAAPFLTYGLFQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 65 summaries

Database :

1: PIR\_73:\*  
 2: PIR1:\*  
 3: PIR2:\*  
 4: PIR3:\*  
 5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	3.5	111	2	A85866
2	9	3.2	733	2	S78376
3	8	2.8	58	2	A58208
4	8	2.8	143	2	G64168
5	8	2.8	197	2	E72374
6	8	2.8	220	2	AG3547
7	8	2.8	278	2	DB3080
8	8	2.8	339	2	C71132
9	8	2.8	372	2	H70813
10	8	2.8	379	2	E64300
11	8	2.8	381	2	AH3041
12	8	2.8	387	2	DB4885
13	8	2.8	397	2	D98244
14	8	2.8	422	1	A60503
15	8	2.8	443	2	TL7220
16	8	2.8	465	2	AC0347
17	8	2.8	471	2	A75267
18	8	2.8	576	2	E64186
19	8	2.8	586	2	A41125
20	7	2.5	45	2	D58208
21	7	2.5	50	2	S22582
22	7	2.5	58	2	S34045
23	7	2.5	86	2	F87604
24	7	2.5	102	2	F87993
25	7	2.5	115	2	PH1560
26	7	2.5	115	2	H83201
27	7	2.5	118	1	IEEC5B
28	7	2.5	118	2	AE1753
29	7	2.5	123	2	AH2707

## ALIGNMENTS

30	7	2.5	125	2	T27519	hypothetical prote
31	7	2.5	131	2	I52290	interleukin-13 - r
32	7	2.5	131	2	E30552	T-cell activation
33	7	2.5	146	2	T37116	probable transpos
34	7	2.5	147	2	A71217	hypothetical prote
35	7	2.5	150	2	T08734	hypothetical prote
36	7	2.5	157	2	S31078	seed allergen RA5
37	7	2.5	157	2	T02664	allergen - rice
38	7	2.5	157	2	A75567	conserved hypotet
39	7	2.5	157	2	E75530	hypothetical prote
40	7	2.5	160	2	S59925	allergen RA5B prec
41	7	2.5	161	1	DNEC17	outer membrane pro
42	7	2.5	161	2	D90651	histone-like prote
43	7	2.5	161	2	D85502	hypothetical prote
44	7	2.5	162	2	T34937	hypothetical prote
45	7	2.5	162	2	T31173	hypothetical prote
46	7	2.5	164	2	S76920	hypothetical prote
47	7	2.5	170	2	S44789	D2007.4 protein -
48	7	2.5	174	2	D87638	transcription regu
49	7	2.5	180	1	LGST	beta-lactoglobulin
50	7	2.5	180	1	LGSH	beta-lactoglobulin
51	7	2.5	181	2	B60738	insulin-like growt
52	7	2.5	187	2	G85343	phospholipase A2-I
53	7	2.5	206	2	S72567	hypothetical prote
54	7	2.5	230	2	AH0692	conserved hypotet
55	7	2.5	231	2	B64920	probable membrane
56	7	2.5	231	2	E90921	hypothetical prote
57	7	2.5	231	2	A85770	hypothetical prote
58	7	2.5	231	2	C86665	amino acid ABC tra
59	7	2.5	233	2	S60767	ribonuclease III -
60	7	2.5	233	2	A80273	probable membrane
61	7	2.5	234	2	G85098	H+-transporting AT
62	7	2.5	235	2	I64174	probable sodium-tr
63	7	2.5	236	2	A86387	probable cytochrom
64	7	2.5	238	2	H70866	hypothetical prote
65	7	2.5	240	2	G83208	conserved hypotet

RESULT 1  
 A85866  
 hypothetical protein Z3516 [imported] - Escherichia coli (strain O157:H7, substrain E  
 C.Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C.Accession: A85866  
 R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A.Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A.Reference number: A85480; MUID:21074935; PMID:11206551  
 A.Accession: A85866  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-111 <STO>  
 A.Cross-references: GB:AE005174; NID:q12516604; PIDN:AMG57389.1; GSPDB:GN00145; UWGP:  
 A.Experimental source: strain O157:H7, substrain EDL933  
 C.Genetics:  
 A.Gene: Z3516

Query Match 3.5%; Score 10; DB 2; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 0.073;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGTALACUGL 72  
 |||||  
 DB 40 LGTALACUGL 49

RESULT 2  
 S78376  
 photosystem I P700 apoprotein A2 - Odontella sinensis chloroplast

C:Species: chloroplast *Odontella sinensis*  
 C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 20-Jun-2000  
 C:Accession: S78376  
 R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroch-Pancic, P.; Frier, U.  
 Plant Mol. Biol. Rep. 13, 336-342, 1995  
 A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, *Odontella sinensis*  
 A:Reference number: S78238  
 A:Accession: S78376  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-733 <KOW>  
 A:Cross-references: EMBL:Z67753; NID:q1185127; PIDN:CA91749.1; PID:q1185266  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995  
 C:Genetics:  
 A:Gene: psab  
 A:Genome: chloroplast  
 C:Superfamily: photosystem I P700 apoprotein  
 C:Keywords: chloroplast; electron transfer; membrane protein; membrane-associated complex

Query Match 3.2%; Score 9; DB 2; Length 733;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGIALACLG 71  
 |||||  
 Db 333 LGIALACLG 341

RESULT 3  
 A58208  
 C:Species: *Chrysemys picta* (painted turtle)  
 C:Date: 08-Nov-1996 #sequence\_revision 08-Nov-1996 #text\_change 07-May-1999  
 C:Accession: A58208  
 R:Hunt, J.G.; Kaslinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.  
 J. Biol. Chem. 271, 23547-23557, 1996  
 A:Title: Proteomes of reptiles.  
 A:Reference number: A58208; MUID:96394458; PMID:8798564  
 A:Accession: A58208  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-58 <HUN>  
 C:Superfamily: sperm histone

Query Match 2.8%; Score 8; DB 2; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 ORRGRRG 49  
 |||||  
 Db 35 ORRGRRG 42

RESULT 4  
 G84168  
 hypothetical protein Vng0080h [imported] - *Halobacterium* sp. NRC-1  
 C:Species: *Halobacterium* sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: G84168  
 R:Ng, W.V.; Kennedy, S.P.; Mahalas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
 Jung, K.H.; Alam, M.; Freites, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of *Halobacterium* species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: G84168  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-143 <STO>  
 A:Cross-references: GB:AE004437; NID:q10579733; PIDN:AG18715.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNG0080H

Query Match 2.8%; Score 8; DB 2; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 7.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
 |||||  
 Db 55 LGIGLALA 62

RESULT 5  
 E72374  
 hypothetical protein - *Thermotoga maritima* (strain MSB8)  
 C:Species: *Thermotoga maritima*  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: E72374  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
 C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
 A:Reference number: A72200; MUID:99287316; PMID:10360571  
 A:Accession: E72374  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-197 <ARN>  
 A:Cross-references: GB:AE001724; GB:AE000512; NID:q4980966; PIDN:AAD35553.1; PID:q498  
 C:Genetics:  
 A:Experimental source: strain MSB8  
 A:Gene: TM0469

Query Match 2.8%; Score 8; DB 2; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 9.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SLIDFEIS 9  
 |||||  
 Db 135 SLIDFEIS 142

RESULT 6  
 AG3547  
 bicyclomycin resistance protein [imported] - *Brucella melitensis* (strain 16M)  
 C:Species: *Brucella melitensis*  
 C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
 C:Accession: AG3547  
 R:DelVecchio, V.G.; Kaparat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov  
 J.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melit*  
 A:Reference number: AD3252; PMID:11756688  
 A:Accession: AG3547  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-220 <KUR>  
 A:Cross-references: GB:AE008918; PIDN:AAL53546.1; PID:q17984455; GSPDB:GN00191  
 A:Experimental source: strain 16M  
 C:Genetics:  
 A:Gene: BMEI10304  
 A:Map position: II

Query Match 2.8%; Score 8; DB 2; Length 220;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 ALLVPLAL 61  
 |||||  
 Db 132 ALLVPLAL 139

RESULT 7  
 D83080  
 hypothetical protein PA4521 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Species: *Pseudomonas aeruginosa*  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: D83080  
 R:Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Micooguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Linn, J.; Lory, S.; Olson, M.V.  
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: D83080  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-278 <STO>  
 A:Cross-references: GB:AE004866; GB:AE004091; NID:g9950760; PIDN:AG07909.1; GSPDB:GN001  
 C:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4521

Query Match 2.8%; Score 8; DB 2; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAVV 77  
 |||||  
 Db 47 LGLLAVV 54

RESULT 8  
 C71132  
 hypothetical protein PH0824 - *Pyrococcus horikoshii*  
 C:Species: *Pyrococcus horikoshii*  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: C71132  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatakeyama, Y.; Hino, Y.; Yamamoto, S.; Sekir M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: C71132  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-339 <RAW>  
 A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BA29917.1; PID:g3257234  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by Genbank  
 C:Genetics:  
 A:Gene: PH0824  
 C:Superfamily: conserved hypothetical protein MTH900

Query Match 2.8%; Score 8; DB 2; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAVV 77  
 |||||  
 Db 111 LGLLAVV 118

RESULT 9  
 H70813  
 probable cysteine synthase - *Mycobacterium tuberculosis* (strain H37RV)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: H70813  
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9654230  
 A:Accession: H70813  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
 A:Residues: 1-372 <COL>  
 A:Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CA17654.1; PID:g291  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: cysM3  
 C:Superfamily: threonine dehydratase

Query Match 2.8%; Score 8; DB 2; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLATA 68  
 |||||  
 Db 96 LGGLATA 103

RESULT 10  
 E64300  
 formate dehydrogenase (EC 1.2.1.2) beta chain - *Methanococcus jannaschii*  
 C:Species: *Methanococcus jannaschii*  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: E64300  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannasc*  
 A:Reference number: A64300; MUID:96337999; PMID:8688087  
 A:Accession: E64300  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-379 <BNL>  
 A:Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AA897986.1; PID:g1590818  
 C:Genetics:  
 A:Map position: REV7250-6111  
 C:Superfamily: formate dehydrogenase chain B; ferredoxin 2[4Fe-4S] homology  
 C:Keywords: oxidoreductase  
 F:273-348/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match 2.8%; Score 8; DB 2; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVYGVLA 224  
 |||||  
 Db 35 LVYGVLA 42

RESULT 11  
 AH3041  
 conserved hypothetical protein Atu3948 [imported] - *Agrobacterium tumefaciens* (strain C:Species: *Agrobacterium tumefaciens*  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C:Accession: AH3041  
 R:Mood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moe erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCl ; Karp, P.; Romero, P.; Zhang, S.  
 A:Title: The genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AH3041  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-381 <KUR>  
 A:Cross-references: GB:AE008689; PIDN:AL44750.1; PID:g17742385; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu3948  
 A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 381;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258  
|||||  
DB 325 LALRPGSS 332

## RESULT 12

D84885

hypothetical protein At2g45000 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C&gt;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: D84885

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beutlo, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;

euss, D.; Niernan, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.;

Nature 402, 761-768, 1999

A&gt;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84885

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-387 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:94895250; PIDN:AA032835.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g45000

A:Map position: 2

Query Match 2.8%; Score 8; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 EELVAED 100  
|||||  
DB 370 EELVAED 377

## RESULT 13

D98244

hypothetical protein AGR\_L1808 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens

C&gt;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C:Accession: D98244

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Mollem, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A&gt;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; PMID:11743194

A:Accession: D98244

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 &lt;KUR&gt;

A:Cross-references: GB:AE007870; PIDN:AAK89478.1; PID:q15159347; GSPDB:GN00170

C:Genetics:

A:Gene: AGR\_L1808

A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 397;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258  
|||||  
DB 341 LALRPGSS 348

## RESULT 14

A60503

sperm-binding glycoprotein ZP3 precursor - golden hamster

N:Alternate names: sperm receptor; zona pellucida glycoprotein ZP3

C:Species: Mesocricetus auratus (golden hamster)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A60503

R:Kinloch, R.A.; Ruiz-Seller, B.; Wasserman, P.M.

Dev. Biol. 142, 414-421, 1990

A&gt;Title: Genomic organization and polypeptide primary structure of zona pellucida gly

A:Reference number: A60503; MUID:91078540; PMID:2257975

A:Accession: A60503

A:Molecule type: DNA

A:Residues: 1-422 &lt;KIN&gt;

A:Cross-references: GB:M63629

A&gt;Note: the authors translated the codon CAA for residue 251 as Glu, and AGC for rest

C:Comment: This suitated glycoprotein in the zona pellucida of the oocyte is a recept

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

C:Keywords: glycoprotein; oocyte

F:45-300/Domain: ZP domain homology &lt;ZPH&gt;

Query Match 2.8%; Score 8; DB 1; Length 422;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLGLA 66  
|||||  
DB 386 LALGLGLA 393

## RESULT 15

T17220

hypothetical protein DKFZ566001.1 - human

C:Species: Homo sapiens (man)

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Jun-2000

C:Accession: T17220

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18725

A:Accession: T17220

A&gt;Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-443 &lt;BLU&gt;

A:Cross-references: EMBL:AL117414

A:Experimental source: fetal kidney; clone DKFZ566001.1

C:Genetics:

A&gt;Note: DKFZ566001.1

C:Superfamily: gamma-glutamyltransferase

Query Match 2.8%; Score 8; DB 2; Length 443;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
|||||  
DB 14 LGIGLALA 21

## RESULT 16

AC0347

probable membrane protein yegB [imported] - Yersinia pestis (strain C092)

C:Species: Yersinia pestis

C&gt;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 17-May-2002

C:Accession: AC0347

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.

deno-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.

11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell

Nature 413, 523-527, 2001

A&gt;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AC0347

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-465 &lt;KUR&gt;

A:Cross-references: GB:AL590842; PIDN:CAC92102.1; PID:q15980820; GSPDB:GN00175

C:Genetics:

A:Gene: yegB

C:Superfamily: multidrug-efflux transporter

Query Match 2.8%; Score 8; DB 2; Length 465;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVSL 79  
DB 335 LLLAVVSL 342

RESULT 17  
A:5267

probable transport protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: A75267

R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mc

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: A75267

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-471 <WHI>

A:Cross-references: GB:AE002079; GB:AE000513; NID:g6460315; PIDN:AAF12043.1; PID:g646032

A:Experimental source: strain R1

A:Genetics:

A:Gene: DR2502

A:Map position: 1

Query Match 2.8%; Score 8; DB 2; Length 471;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLAL 67  
DB 366 ALGLGLAL 373

RESULT 18  
E64186

probable ATP-binding transport protein H1156 - *Haemophilus influenzae* (strain Rd KW20)

C:Species: *Haemophilus influenzae*

C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 02-Feb-2001

C:Accession: E64186

R.Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: E64186

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-576 <TIGR>

A:Cross-references: GB:U32795; GB:I42023; NID:91574708; PIDN:AAC22811.1; PID:91574713;

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop

F:355-550/Domain: ATP-binding cassette homology <ABC>

F:372-379/Region: nucleotide-binding motif A (P-loop)

Query Match 2.8%; Score 8; DB 2; Length 576;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 58 PLALGLGL 65  
DB 159 PLALGLGL 166

RESULT 19  
A41125

gamma-glutamyltransferase (EC 2.3.2.2) related protein - human

N/Alternate names: gamma-glutamyltransferase-like activity 1; GGT-REL

C:Species: *Homo sapiens* (man)

C>Date: 27-Mar-1992 #sequence\_revision 27-Mar-1992 #text\_change 18-Jun-1999

C:Accession: A41125

R.Heisterkamp, N.; Rajpert-De Meyts, E.; Uribe, L.; Forman, H.J.; Groffen, J.

Proc. Natl. Acad. Sci. U.S.A. 88, 6303-6307, 1991

A:Title: Identification of a human gamma-glutamyl cleaving enzyme related to, but dis

A:Reference number: A41125; MUID:91296809; PMID:1676842

A:Accession: A41125

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-586 <HEI>

A:Cross-references: GB:M64099; NID:9183141; PIDN:AAA58503.1; PID:9183142

C:Genetics:

A:Gene: GDB:GOTL1; GGT-REL

A:Cross-references: GDB:134033

C:Superfamily: gamma-glutamyltransferase

C:Keywords: aminocyltransferase; glycoprotein; transmembrane protein

Query Match 2.8%; Score 8; DB 2; Length 586;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLGLA 68  
DB 14 LGGLGLA 21

RESULT 20  
D58208

protamine II-3 - painted turtle

C:Species: *Chrysemys picta* (painted turtle)

C>Date: 08-Nov-1996 #sequence\_revision 08-Nov-1996 #text\_change 07-May-1999

C:Accession: D58208

R.Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp,

J. Biol. Chem. 271, 23547-23557, 1996

A:Title: Protamines of reptiles.

A:Reference number: A58208; MUID:96394458; PMID:8798564

A:Accession: D58208

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-45 <HUN>

C:Superfamily: sperm histone

Query Match 2.5%; Score 7; DB 2; Length 45;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
DB 24 RRRGRG 30

RESULT 21  
S22582

protamine 1 - *Saguinus imperator*

C:Species: *Saguinus imperator*

C>Date: 22-Nov-1993 #sequence\_revision 12-Apr-1996 #text\_change 21-Jul-2000

C:Accession: S22582

R.Queral, R.; Oliva, R.

Nucleic Acids Res. 19, 5786, 1991

A:Title: Protamine 1 gene sequence from the primate *Saguinus imperator* isolated with

A:Reference number: S22582; MUID:92051332; PMID:1840669

A:Accession: S22582

A:Molecule type: DNA

A:Residues: 1-50 <QUE>

A:Cross-references: EMBL:X61678; NID:958405; PIDN:CAA3853.1; PID:94494091

A>Note: the authors translated the codon TAC for residue 43 as Thr

C:Genetics:

A:Introns: 37/1  
 C:Superfamily: sperm histone  
 C:Keywords: chromosomal protein; DNA binding; nucleus; spermatogenesis

Query Match 2.5%; Score 7; DB 2; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 QRRGR 48  
 |||||  
 Db 18 QRRGR 24

RESULT 22

S34045  
 protamine - North American opossum  
 C:Species: Didelphis virginiana, Didelphis marsupialis virginiana (North American opossum)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999

C:Accession: S34045  
 R:Winkler, R.J.; Nishikawa, S.; Connor, W.; Dixon, G.H.

A:Title: Characterization of a marsupial sperm protamine gene and its transcripts from t  
 A:Reference number: S34045; MUID:93345500; PMID:8344286

A:Accession: S34045  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-58 <WIN>

A:Cross-references: EMBL:X74044; NID:g407062; PIDN:CAA52193.1; PID:g407063

C:Superfamily: sperm histone  
 C:Keywords: DNA binding; nucleus

Query Match 2.5%; Score 7; DB 2; Length 58;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 |||||  
 Db 35 RRRGRG 41

RESULT 23

F87604

hypothetical protein CC2870 [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: F87604  
 R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete genome sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: F87604  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-86 <STO>

A:Cross-references: GB:AE005673; NID:g13424486; PIDN:AAK24834.1; GSPDB:GN00148

C:Genetics:  
 A:Gene: CC2870

Query Match 2.5%; Score 7; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLGL 65  
 |||||  
 Db 26 LALGLGL 32

RESULT 24

F87993  
 protein ZC334.3 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: F87993  
 R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;

A:Accession: F87993  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-102 <STO>

A:Cross-references: GB:chr\_I; PIDN:CA804964.1; PID:g3881432; GSPDB:GN00019; CESP:ZC33

A:Note: predicted using GeneFinder  
 C:Genetics:  
 A:Gene: ZC334.3

A:Map position: 1  
 Query Match 2.5%; Score 7; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LVPPLAL 61  
 |||||  
 Db 33 LVPPLAL 39

RESULT 25

PH1560

Ig heavy chain V region (clone VH32) - human (fragment)  
 C:Species: Homo sapiens (man)

C:Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 21-Jan-2000

C:Accession: PH1560  
 R:Rassenti, L.Z.; Kipps, T.J.

J. Exp. Med. 177, 1039-1046, 1993  
 A:Title: Lack of extensive mutations in the VH5 genes used in common B cell chronic 1

A:Reference number: PH157; MUID:93210459; PMID:7681468

A:Accession: PH1560  
 A:Molecule type: DNA  
 A:Residues: 1-115 <RAS>

C:Superfamily: immunoglobulin V region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin

F:32-115/Domain: immunoglobulin homology <IMM>

Query Match 2.5%; Score 7; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAV 76  
 |||||  
 Db 5 LGLLAV 11

RESULT 26

H83201

conserved hypothetical protein PA3557 [imported] - Pseudomonas aeruginosa (strain PAO  
 C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: H83201  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L

.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10964043

A:Accession: H83201  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-115 <STO>

A:Cross-references: GB:AE004776; GB:AE004091; NID:g9949701; PIDN:AA606945.1; GSPDB:GN  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA3557



Query Match 2.5%; Score 7; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGIGL 65  
 |||||  
 Db 49 LALGIGL 55

## RESULT 27

IEEC5B  
 hypothetical protein, 12K - Escherichia coli insertion sequence ISS

C:Species: Escherichia coli  
 C:Date: 18-Dec-1981 #sequence\_revision 18-Dec-1981 #text\_change 10-Sep-1999  
 C:Accession: B91483; C03582; A04466

R:Schoner, B.; Kahn, M.  
 Gene 14, 165-174, 1981

A:Title: The nucleotide sequence of ISS from Escherichia coli.

A:Reference number: A91483; MUID:82028653; PMID:6269959

A:Accession: B91483

A:Molecule type: DNA

A:Residues: 1-118 <SCH>

R:Engler, J.A.; van Bree, M.P.

Gene 14, 155-163, 1981

A:Title: The nucleotide sequence and protein-coding capability of the transposable element

A:Reference number: A03582; MUID:82028652; PMID:6269958

A:Accession: C03582

A:Molecule type: DNA

A:Residues: 1-118 <ENG>

C:Genetics:

A:Mobile element: insertion sequence ISS

C:Superfamily: Escherichia coli hypothetical protein o263

Query Match 2.5%; Score 7; DB 1; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 DGAVRQ 29  
 |||||  
 Db 107 DGAVRQ 113

## RESULT 28

AE1753

Orf1 (bacteriophage b1285) homolog lin2570 [imported] - Listeria innocua (strain Clp1)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C:Accession: AE1753

R:Glaser, P.; Franzen, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

Science 294, 849-852, 2001

A:Authors: Kieft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madheno, E.; Maitournam, A.; Ma

ok, C.; Schluter, T.; Simoes, N.; Tlerraz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AE1753

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-118 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97797.1; PID:916415092; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lin2570

Query Match 2.5%; Score 7; DB 2; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SLDFEI 8  
 |||||  
 Db 61 SLDFEI 67

## RESULT 29

AH2707

conserved hypothetical protein Atu1065 [imported] - Agrobacterium tumefaciens (strain

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002

C:Accession: AH2707

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl

ard, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AH2707

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-123 <KUR>

A:Cross-references: GB:AE008688; PIDN:AA142078.1; PID:917739458; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu1065

A:Map position: circular chromosome

Query Match 2.5%; Score 7; DB 2; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 RLPLPRS 18  
 |||||  
 Db 50 RLPLPRS 56

## RESULT 30

T27519

hypothetical protein ZC334.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T27519

R:McLay, K.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20381

A:Accession: T27519

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-125 <WIL>

A:Cross-references: EMBL:Z82082; PIDN:CA804964.2; GSPDB:GN00019; CESP:ZC334.3

A:Experimental source: clone ZC334

C:Genetics:

A:Gene: CESP:ZC334.3

A:Map position: 1

A:introns: 99/1

Query Match 2.5%; Score 7; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LVLPLAL 61  
 |||||  
 Db 33 LVLPLAL 39

## RESULT 31

I52290

Interleukin-13 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999

C:Accession: I52290

R:Lakkis, F.G.; Cruet, E.N.

Biochem. Biophys. Res. Commun. 197, 612-618, 1993

A:Title: Cloning of rat Interleukin-13 (IL-13) cDNA and analysis of IL-13 gene expres

A:Reference number: I52290; MUID:94092138; PMID:7916615  
 A:Accession: I52290  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-131 <RES>  
 A:Cross-references: GB:U26913; NID:9438875; PIDN:AAA16478.1; PID:9438876  
 C:Genetics:  
 A:Gene: IL-13  
 C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 71  
 |||||  
 DB 9 LALACLG 15

## RESULT 32

E30552

T-cell activation protein p60 precursor - mouse

C:Species: Mus musculus (house mouse)

C&gt;Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 16-Jul-1999

C:Accession: E30552

R:Brown, K.D.; Zurawski, S.M.; Mosmann, T.R.; Zurawski, G.

J. Immunol. 142, 679-687, 1989

A:Title: A family of small inducible proteins secreted by leukocytes are members of a new class of various activation processes.

A:Reference number: A30552; MUID:89093958; PMID:2521553

A:Accession: E30552

A:Molecule type: mRNA

A:Residues: 1-131 &lt;BRO&gt;

A:Cross-references: GB:M23504; NID:9533246; PIDN:AAA40149.1; PID:9533247

C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 71  
 |||||  
 DB 9 LALACLG 15

## RESULT 33

T37116

probable transposase, truncated [imported] - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C&gt;Date: 08-Sep-2000 #sequence\_revision 08-Sep-2000 #text\_change 15-Sep-2000

C:Accession: T37116

R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21588

A:Accession: T37116

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-146 &lt;SAU&gt;

A:Cross-references: EMBL:AL109950; PIDN:GAB52967.1; GSPDB:GN00070; SCOEDB:SCJ4.33c

A:Experimental source: strain A312

C:Genetics:  
 A:Gene: SCOEDB:SCJ4.33c

C:Superfamily: Synchocystis transposase s111710

Query Match 2.5%; Score 7; DB 2; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAV 76  
 |||||  
 DB 102 LGLLAV 108

## RESULT 34

A71217

hypothetical protein PH2001 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C&gt;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000

C:Accession: A71217

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushiida, N.; Ogu

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: A71217

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-147 &lt;KAW&gt;

A:Cross-references: GB:AP000007; GB:AP000001; NID:93236134; NID:93236128; PID:9325844

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:  
 A:Gene: PH2001

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH2001

Query Match 2.5%; Score 7; DB 2; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 LACIGLL 73  
 |||||  
 DB 45 LACIGLL 51

## RESULT 35

T08734

hypothetical protein DKFZp566F0546.1 - human (fragment)

C:Species: Homo sapiens (man)

C&gt;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 13-Aug-1999

C:Accession: T08734

R:Ottensmeyer, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16474

A:Accession: T08734

A:Molecule type: mRNA

A:Residues: 1-150 &lt;OTT&gt;

A:Cross-references: EMBL:AL050075

A:Experimental source: fetal kidney; clone DKFZp566F0546

C:Genetics:  
 A:Note: DKFZp566F0546.1

Query Match 2.5%; Score 7; DB 2; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 234 AASSLGP 240  
 |||||  
 DB 32 AASSLGP 38

## RESULT 36

S31078

seed allergen RA5 - rice

C:Species: Oryza sativa (rice)

C&gt;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000

C:Accession: S31078

R:Adachi, T.; Izumi, H.; Yamada, T.; Tanaka, K.; Takeuchi, S.; Nakamura, R.; Matsuda,

Plant Mol. Biol. 21, 239-248, 1993

A:Title: Gene structure and expression of rice seed allergenic proteins belonging to

A:Reference number: S31078; MUID:93144699; PMID:7678765

A:Accession: S31078

A:Molecule type: mRNA

A:Residues: 1-157 &lt;ADA&gt;

A:Cross-references: EMBL:D11330; NID:g218196; PIDN:BA01996.1; PID:g218197

C:Superfamily: wheat alpha-amylase inhibitor

C:Keywords: seed

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78  
|||||  
DB 11 LLLAVS 17

RESULT 37  
T02664

allergen - rice

C:Species: Oryza sativa (rice)

C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Jul-1999

C:Accession: T02664

R:Yun, C.H.; Park, J.H.; Eun, M.Y.

submitted to the EMBL Data Library, January 1998

A:Description: Nucleotide sequence of rice allergenic protein.

A:Reference number: 214691

A:Accession: T02664

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-157 <YUN>

A:Cross-references: EMBL:AF042200; NID:g2827315; PIDN:AAB99797.1; PID:g2827316

A:Experimental source: strain Nipponbare

C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78  
|||||  
DB 11 LLLAVS 17

RESULT 38  
A75567

conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C:Accession: A75567

R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: A75567

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <WHI>

A:Cross-references: GB:AE001867; GB:AE000513; NID:g6457693; PIDN:AAF09626.1; PID:g6457693

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0033

A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 IRTLPMA 267  
|||||  
DB 22 IRTLPMA 28

RESULT 39  
E75530

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: E75530

R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75530

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <WHI>

A:Cross-references: GB:AE001895; GB:AE000513; NID:g6458024; PIDN:AAF09935.1; PID:g645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0352

A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRGRRG 49  
|||||  
DB 144 RRGRRG 150

RESULT 40  
S59925

allergen RASB precursor - rice

C:Species: Oryza sativa (rice)

C>Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Jun-2000

C:Accession: S59925

R:Alvarez, A.M.; Adachi, T.; Nakase, M.; Aoki, N.; Nakamura, R.; Matsuda, T.

Biochim. Biophys. Acta 1251, 201-204, 1995

A:Title: Classification of rice allergenic protein cDNAs belonging to the alpha-amylase

A:Reference number: S59922; MUID:95399441; PMID:7669811

A:Accession: S59925

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-160 <ALV>

A:Cross-references: EMBL:D42142; NID:g1398917; PIDN:BAA07713.1; PID:g1398918

C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78  
|||||  
DB 11 LLLAVS 17

Search completed: March 31, 2003, 14:16:42  
Job time : 47 secs

**THIS PAGE BLANK (USPTO)**

Brue  
RS 5/15/03  
Page 1

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 11:55:46 ; Search time 22 Seconds  
(without alignments)  
535,422 Million cell updates/sec

Title: US-09-245-198A-4  
Perfect score: 284  
Sequence: 1 MSLLDFEISARRLPRLPSLG.....PMAHLKAAPFLTFGLFQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 65 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	87.7	249	1	TN12_HUMAN
2	32	11.3	225	1	TN12_MOUSE
3	10	3.5	111	1	YFBW_ECOLI
4	9	3.2	733	1	PSAB_ODOSI
5	9	3.2	734	1	PSAB_CVACA
6	8	2.8	179	1	ADHS_GLUOX
7	8	2.8	220	1	Y304_BRUME
8	8	2.8	317	1	MSHR_PANTR
9	8	2.8	379	1	FDHB_METUA
10	8	2.8	422	1	2P3_MESAV
11	8	2.8	576	1	CYDC_HAEN
12	8	2.8	586	1	GCT5_HUMAN
13	7	2.5	49	1	HSP1_SAGIM
14	7	2.5	57	1	HSP1_DIDMA
15	7	2.5	115	1	A62F_DROME
16	7	2.5	118	1	Y151_ECOLI
17	7	2.5	131	1	IL13_MOUSE
18	7	2.5	131	1	IL13_RAT
19	7	2.5	147	1	YK01_PYRHO
20	7	2.5	150	1	TNFC_PIG
21	7	2.5	157	1	RA05_ORYSA
22	7	2.5	161	1	HUPA_ECOLI
23	7	2.5	170	1	YLMA_CAEEL
24	7	2.5	180	1	LACB_BUBBU
25	7	2.5	180	1	LACB_CAPHI
26	7	2.5	180	1	LACB_SHEEP
27	7	2.5	230	1	RNFE_SALTY
28	7	2.5	231	1	RNFE_ECO57
29	7	2.5	231	1	RNFE_ECOLI
30	7	2.5	233	1	RNFE_COXBU
31	7	2.5	233	1	RNFE_YERPE
32	7	2.5	235	1	RNFE_HAEIN
33	7	2.5	239	1	TN14_MOUSE

## ALIGNMENTS

34	7	2.5	240	1	RNFE_PSEAE	09hyb5 pseudomonas
35	7	2.5	243	1	MOEB_HAEIN	P45211 haemophilus
36	7	2.5	249	1	TNFC_HUMAN	006643 homo sapien
37	7	2.5	249	1	MOEB_ECOLI	P12282 escherichia
38	7	2.5	249	1	MOEB_SALTY	056057 salmonella
39	7	2.5	257	1	KDTX_SERMA	054435 serattia ma
40	7	2.5	310	1	TNFC_MARMO	09jml0 marmota mon
41	7	2.5	316	1	ISPH_XYLFA	09pas9 xyliella fas
42	7	2.5	324	1	ISPH_XYLFA	P21882 bacillus su
43	7	2.5	329	1	ODPB_BACSU	009208 caenorhabdi
44	7	2.5	335	1	SRA6_CAEEL	035777 rattus norv
45	7	2.5	344	1	LEU3_THEAO	P24098 thermus aqu
46	7	2.5	357	1	G6PT_MOUSE	P35576 mus musculu
47	7	2.5	357	1	G6PT_MOUSE	P43428 rattus norv
48	7	2.5	358	1	PONT_RABIT	P27170 oryctolagus
49	7	2.5	365	1	NOO8_THETH	060019 thermus the
50	7	2.5	394	1	BENE_ACICA	P07775 acinetobact
51	7	2.5	396	1	DH11_XENLA	091610 xenopus lae
52	7	2.5	398	1	DH12_XENLA	091611 xenopus lae
53	7	2.5	402	1	SELP_BOVIN	P49907 bos taurus
54	7	2.5	412	1	PKRP_ALCEU	P50320 alcaligenes
55	7	2.5	413	1	PKRC_ALCEU	P50319 alcaligenes
56	7	2.5	416	1	NH59_CAEEL	09txj1 caenorhabdi
57	7	2.5	418	1	CP16_RAT	P09006 rattus norv
58	7	2.5	419	1	ENO_PYRAE	08zy67 pyrobaculum
59	7	2.5	423	1	YJ54_YEAST	P47130 saccharomyc
60	7	2.5	424	1	TP3_MOUSE	P10761 mus musculu
61	7	2.5	428	1	SYH_CHLMU	09pj19 chilamydia m
62	7	2.5	455	1	PHR_STRGR	P12768 streptomyce
63	7	2.5	461	1	PUCG_RHOCA	P23462 rhodobacter
64	7	2.5	461	1	Y608_HAEIN	057486 haemophilus
65	7	2.5	464	1	SOXB_MOUSE	004866 mus musculu

## ALIGNMENTS

1	TN12_HUMAN	STANDARD:	PRT:	249 AA.
ID	TN12_HUMAN			
AC	043508; Q8WU27;			
DT	15-JUN-2002 (Rel. 41, Created)			
DR	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).			
GN	TNFSF12 OR APO3L OR DR3LG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. AND N-TERMINUS OF SOLUBLE FORM.			
RC	TISSUE=Tonsil, and Fetal Liver;			
RX	MEDLINE=98070415; PubMed=9405449;			
RA	Chichportliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,			
RT	Hession C., Garcia I., Browning J.L.,			
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."			
RL	J. Biol. Chem. 272:32401-32410(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal kidney;			
RX	MEDLINE=98228355; PubMed=9560343;			
RA	Marsters S.A., Sheridan J.P., Pilti R.M., Brush J., Goddard A.,			
RT	Ashkenazi A.;			
RT	"Identification of a ligand for the death-domain-containing receptor Apo3".			
RL	Curr. Biol. 8:525-528(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Tonsil;			
RA	Strausberg R.;			

```

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION.
RA Pubmed=10085077;
RX Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;
RT "TWEAK induces angiogenesis and proliferation of endothelial cells.";
RL J. Biol. Chem. 274:8455-8459(1999).
CC -1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/AP03. Weak
CC inducer of apoptosis in some cell types. Mediates NF-kappaB
CC activation. May promote angiogenesis and the proliferation of
CC endothelial cells.
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.
CC -1- TISSUE SPECIFICITY: Highly expressed in adult heart, pancreas,
CC skeletal muscle, brain, colon, small intestine, lung, ovary,
CC prostate, spleen, lymph node, appendix and peripheral blood
CC lymphocytes. Low expression in kidney, testis, liver, placenta,
CC thymus and bone marrow. Also detected in fetal kidney, liver,
CC lung and brain.
CC -1- PTM: The soluble form derives from the membrane form
CC by proteolytic processing.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift in position 125.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF030099; AAC51923.1; -
DR EMBL: AF055872; AAC39724.1; -
DR EMBL: BC019047; AAH19047.1; ALT_FRAME.
DR Genew; HGNC:11927; TNFSF12.
DR MIM: 602695; -
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; FALSE_NEG.
DR PROSITE: PS50049; TNF_2; 1.
KW Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
FT FT CHAIN 1 249
FT FT CHAIN 1 249
FT FT CHAIN 94 249
FT FT DOMAIN 1 21
FT FT TRANSMEM 22 42
FT FT TRANSMEM 22 42
FT FT DOMAIN 1 21
FT FT SITE 43 249
FT FT CARBOHYD 139 139
FT FT SEQUENCE 249 AA; 27216 MW; E66084361C28BBA CRC64;
SO QUERY MATCH 87.7%; Score 249; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 3.2e-231;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 181 LTVDSVLAIRCEEFSSATASSLGPOLRLCQVSGLLALRPGSSLRITRLPMHLKAPFL 240
OY 276 TYEGLFQVH 284
DB 241 TYEGLFQVH 249
-----
RESULT 2
TNF2_MOUSE STANDARD: PRT: 225 AA.
ID TNF2_MOUSE
AC 054907: 09CTP2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak
DE inducer of apoptosis) (TWEAK) (Fragment).
GN TNFSF12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal macrophage;
RX MEDLINE=98070415; Pubmed=9405449;
RA Chicheportliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RA Hession C., Garcia I., Browning J.L.;
RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that
RT weakly induces apoptosis.";
RL J. Biol. Chem. 272:32401-32410(1997).
RN [2]
RP SEQUENCE OF 83-225 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl I., Suzuki R., Tomita M., Wagner L., Wasth T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Williams L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/AP03. Weak
CC inducer of apoptosis in some cell types. Promotes angiogenesis and
CC the proliferation of endothelial cells. Mediates NF-kappaB
CC activation (By similarity).
CC -1- SUBUNIT: Homotrimer (Potential).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- PTM: The soluble form is produced from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

DR EMBL: AF030100; AAC53517.1; -  
DR EMBL: AK020909; BAB32249.1; -  
DR MGD: MG1:1196259; Tnfst12.  
DR InterPro: IPR000478; TNF\_family.  
DR Pfam: PF00229; TNF; 1.  
DR SMART: SM00207; TNF; 1.  
DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
DR PROSITE: PS50049; TNF\_2; 1.  
DR Cyclokin: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
FT CHAIN 1 1  
FT NON\_TER 1  
FT CHAIN 1 1  
FT CHAIN 70 225  
FT TRANSMEM <1 21  
FT DOMAIN 22 225  
FT SITE 69 70  
FT DISULFID 167 186  
FT CARBOHYD 115 115  
SQ SEQUENCE 225 AA; 24781 MW; 90C412CC0480659B CRC64;

Query Match 11.3%; Score 32; DB 1; Length 225;  
Best Local Similarity 100.0%; Pred. No. 4,7e-23;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 RRAIAHYEHVHPGQDGAQAGVDGTVSGWEE 170  
DB 80 RRAIAHYEHVHPGQDGAQAGVDGTVSGWEE 111

RESULT 3  
YFBW\_ECOLI STANDARD; PRT; 111 AA.  
AC 047377;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein yfbw.  
GN YFBW OR B2257.1.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RA MEDLINE=97426617; Pubmed=9278503;  
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
RA "The complete genome sequence of Escherichia coli K-12";  
RL [2]  
RN [2]  
RP SEQUENCE OF 81-111 FROM N.A.  
RC STRAIN=K12;  
RA MEDLINE=96186953; Pubmed=8626063;  
RA Sharma V., Hudspeh M.E., Meganathan R.;  
RT "Menaquinone (vitamin K2) biosynthesis: localization and characterization of the menE gene from Escherichia coli.";  
RL Gene 168.43-48(1996).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

DR EMBL: AE000315; -; NOT\_ANNOTATED\_CDS.  
DR EMBL: L35031; AAB04895.1; -  
DR Ecogene: EG14344; yfbw.  
DR InterPro: IPR000620; DUF6.  
DR Pfam: PF00892; DUF6; 1.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 36 56  
FT TRANSMEM 62 82  
FT TRANSMEM 88 108  
SQ SEQUENCE 111 AA; 12192 MW; 7CFA06D75DA33D69 CRC64;

Query Match 3.5%; Score 10; DB 1; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.033;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGLALACGL 72  
DB 40 LGLALACGL 49

RESULT 4  
PSAB\_ODOSI STANDARD; PRT; 733 AA.  
AC P49480;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Photosystem I P700 chlorophyll A apoprotein A2 (Psab) (PSI-B).  
GN PSAB.  
OS Odontella sinensis (Marine centric diatom).  
OC Chloroplast.  
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae; Bidulophyceidae; Eupodiscales; Eupodisaceae; Odontella.  
OX NCBI\_TaxID=2839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kowalik K.V., Stoebe B., Schafran I., Kroth-Pancic P., Freier U.;  
RT "The chloroplast genome of a chlorophyll a-bc-containing alga, Odontella sinensis";  
RL Plant Mol. Biol. Rep. 13:336-342(1995).  
CC -1- FUNCTION: Psab and psab bind P700, the primary electron donor of photosystem I (PSI), as well as the electron acceptors A0, A1, and Fx. PSI functions as a plastocyanin/cytochrome c6-ferredoxin oxidoreductase.  
CC -1- COFACTOR: P700 is a chlorophyll A dimer. A0 is chlorophyll A, A1 is a phytylquinone and Fx is a 4Fe-4S iron-sulfur center.  
CC -1- SUBUNIT: A psaa/b heterodimer binds the P700 chlorophyll spectral pair and subsequent electron acceptors. The PSI reaction center of higher plants and algae is composed of one at least 11 subunits.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast thylakoid membrane.  
CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

DR EMBL: Z67753; CAA91749.1; -  
DR HSSP: P25897; L180.  
DR InterPro: IPR001280; PSI\_Psaa/B.  
DR Pfam: PF00223; psaa-psab; 1.  
DR PRINTS: PR00257; PHOTOSPSAB.  
DR PROSITE: PS00419; PHOTOSYSTEM\_I\_PSAAB; 1.  
KW Photosynthesis; Photosystem I; Transport; Electron transport; Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;  
KW Iron-sulfur; 4Fe-4S; Chlorophyll.  
FT TRANSMEM 46 69  
FT TRANSMEM 134 157  
II (POTENTIAL).

```

FT TRANSMEM 174 198 III (POTENTIAL).
FT TRANSMEM 272 290 IV (POTENTIAL).
FT TRANSMEM 329 352 V (POTENTIAL).
FT TRANSMEM 368 394 VI (POTENTIAL).
FT TRANSMEM 416 438 VII (POTENTIAL).
FT TRANSMEM 516 534 VIII (POTENTIAL).
FT TRANSMEM 574 595 IX (POTENTIAL).
FT TRANSMEM 642 664 X (POTENTIAL).
FT TRANSMEM 706 726 XI (POTENTIAL).
FT METAL 558 558 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 567 567 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT BINDING 653 653 P700 SPECIAL PAIR CHLOROPHYLL AXIAL
FT BINDING 661 661 LIGAND (BY SIMILARITY).
FT BINDING 669 669 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 670 670 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 699 699 A1 PHYLOQUINONE (BY SIMILARITY).
FT BINDING 699 699 A1 PHYLOQUINONE (BY SIMILARITY).
SQ SEQUENCE 733 AA; 82103 MW; 13439AF1E41BEF7 CRC64;

Query Match 3.2%; Score 9; DB 1; Length 733;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGIALACLG 71
DB 333 LGIALACLG 341

RESULT 5
PSAB_CYACA STANDARD; PRT: 734 AA.
ID PSAB_CYACA
AC 09TLO6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I P700 chlorophyll A apoprotein A2 (Psab) (PSI-B).
GN PSAB.
OS Cyanidium caldarium.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Cyanidium.
OX NCBI_TaxID=2771;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-RK-1;
RX MEDLINE=20496959; PubMed=11040290;
RA Gloeckner G., Rosenthal A., Valentin K.-U.;
RT "The structure and gene repertoire of an ancient red algal plastid genome.";
RT J. Mol. Evol. 51:382-390(2000).
RL J. Mol. Evol. 51:382-390(2000).
CC -I- FUNCTION: Psaa and psab bind P700, the primary electron donor of photosystem I (PSI), as well as the electron acceptors A0, A1, and FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin oxidoreductase.
CC -I- COFACTOR: P700 is a chlorophyll a dimer, A0 is chlorophyll A, A1 is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
CC -I- SUBUNIT: A psaa/b heterodimer binds the P700 chlorophyll special pair and subsequent electron acceptors. The PSI reaction center of higher plants and algae is composed of one at least 11 subunits.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast thylakoid membrane.
CC -I- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF023186; AAF12881.1; -
CC HSSP: P25897; IJBO.

```

```

DR InterPro: IPR001280; PSI_Psaa/B.
DR Pfam: PF00223; psaa-psab; 1.
DR PRINTS: PR00257; PHOTOSYSAB.
DR PROSITE: PS00419; PHOTOSYSTEM_I_PSAAB; 1.
KM Photosynthesis; Photosystem I; Transport; Electron transport;
KM Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;
KW Iron-sulfur; 4Fe-4S; Chlorophyll.
FT TRANSMEM 46 69 I (POTENTIAL).
FT TRANSMEM 135 158 II (POTENTIAL).
FT TRANSMEM 175 199 III (POTENTIAL).
FT TRANSMEM 273 291 IV (POTENTIAL).
FT TRANSMEM 330 353 V (POTENTIAL).
FT TRANSMEM 369 395 VI (POTENTIAL).
FT TRANSMEM 417 439 VII (POTENTIAL).
FT TRANSMEM 517 535 VIII (POTENTIAL).
FT TRANSMEM 575 596 IX (POTENTIAL).
FT TRANSMEM 643 665 X (POTENTIAL).
FT TRANSMEM 707 727 XI (POTENTIAL).
FT METAL 559 559 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 568 568 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT BINDING 654 654 P700 SPECIAL PAIR CHLOROPHYLL AXIAL
FT BINDING 662 662 LIGAND (BY SIMILARITY).
FT BINDING 670 670 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 671 671 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 700 700 A1 PHYLOQUINONE (BY SIMILARITY).
FT BINDING 700 700 A1 PHYLOQUINONE (BY SIMILARITY).
SQ SEQUENCE 734 AA; 82359 MW; 4496A2AE59CA9B9 CRC64;

Query Match 3.2%; Score 9; DB 1; Length 734;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGIALACLG 71
DB 334 LGIALACLG 342

RESULT 6
ADHS_GLUOX STANDARD; PRT: 179 AA.
ID ADHS_GLUOX
AC 005544;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alcohol dehydrogenase 15 kDa subunit precursor (G3-ADH subunit III).
GN ADHS.
OS Gluconobacter oxydans (Gluconobacter suboxydans).
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconobacter.
OX NCBI_TaxID=442;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-40.
RP STRAIN-IFO 12528;
RX MEDLINE=97208225; PubMed=9055427;
RA Kondo K., Horiouchi S.;
RT "Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their RT expression in Acetobacter pasteurianus.";
RL Appl. Environ. Microbiol. 63:1131-1138(1997).
CC -I- FUNCTION: NOT ESSENTIAL FOR ALCOHOL DEHYDROGENASE ACTIVITY.
CC -I- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -I- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC

```



DR EMBL; D86440; BAA19756.1; -  
KM Membrane: Periplasmic: Signal.  
FT SIGNAL 1 24 POTENTIAL.  
FT CHAIN 25 179 ALCOHOL DEHYDROGENASE 15 KDA SUBUNIT.  
FT MOD\_RES 25 25 PYRROLIDONE CARBOXYLIC ACID.  
SQ SEQUENCE 179 AA; 19943 MW; P6AF24365B3C66 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 179;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGIGLA 66  
|||||  
DB 11 LALGIGLA 18

RESULT 7  
Y304\_BRUME STANDARD; PRT: 220 AA.  
ID Y304\_BRUME  
AC Q8YD73;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein BME110304.  
GN BME110304.  
OS Brucella melitensis.  
OC Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group:  
OC Brucellaceae: Brucella.  
OX NCBI\_TaxID=29459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=16M / ATCC 23456 / Biotype 1;  
RX MEDLINE=20020109; PubMed=11756688;  
RA Delvecchio V.G., Kapratel V., Redkar R.J., Patra G., Mujer C., Los T.,  
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,  
RA Jabolonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,  
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-D.,  
RA Haselkorn R., Kyrildes N., Overbeek R.;  
RT "The genome sequence of the facultative intracellular pathogen  
RT Brucella melitensis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE ODF0191 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE009669; AAL53546.1; -  
KM Hypothetical protein; Transmembrane: Complete proteome.  
FT TRANSMEM 20 39 POTENTIAL.  
FT TRANSMEM 54 72 POTENTIAL.  
FT TRANSMEM 85 104 POTENTIAL.  
FT TRANSMEM 124 146 POTENTIAL.  
FT TRANSMEM 153 175 POTENTIAL.  
FT TRANSMEM 179 198 POTENTIAL.  
SQ SEQUENCE 220 AA; 24815 MW; 182C0244743B17A CRC64;

Query Match 2.8%; Score 8; DB 1; Length 220;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 ALVPLAL 61  
|||||  
DB 132 ALVPLAL 139

RESULT 8  
MSHR\_PANTR

ID MSHR\_PANTR STANDARD; PRT: 317 AA.  
AC O9NFK4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Melanocyte stimulating hormone receptor (MSH-R) (Melanotropin  
DE receptor) (Melanocortin-1 receptor) (MCL-R).  
GN MCLR.  
OS Pan troglodytes (Chimpanzee).  
OC Eumetazoa: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
OC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Pan.  
OX NCBI\_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rees J.L., Harding R.M., Healy E., Jackson I.J., Ray A.J., Ellis N.S.,  
RA Flanagan N., Todd C., Dixon C., Matthews J.N., Sajanilla A.,  
RA Birch-Machin M.A.;  
RT "Chimpanzee melanocortin 1 sequence.";  
CC Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: RECEPTOR FOR MSH (ALPHA, BETA AND GAMMA) AND ACTH. THE  
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE  
CC ADENYLYLATE CYCLASE (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AJ245705; CAB53398.1; -  
DR InterPro: IPR00276; GPCR\_Rhodpsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCR\_Rhodopsn.  
DR PROSITE: PS00237; G-PROTEIN\_RECPT\_FL1; 1.  
DR PROSITE: PS0262; G-PROTEIN\_RECPT\_FL2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 38 63 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 64 72 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 73 93 2 (POTENTIAL).  
FT DOMAIN 94 118 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 119 140 3 (POTENTIAL).  
FT DOMAIN 141 163 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 164 183 4 (POTENTIAL).  
FT DOMAIN 184 191 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 192 211 5 (POTENTIAL).  
FT DOMAIN 212 240 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 241 266 6 (POTENTIAL).  
FT DOMAIN 267 279 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 280 300 7 (POTENTIAL).  
FT DOMAIN 301 317 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 29 29 N-LINKED (GLCNAc...) (POTENTIAL).  
FT LIPID 315 315 PALMITATE (POTENTIAL).  
SQ SEQUENCE 317 AA; 34699 MW; 6615D214E1D247F CRC64;

Query Match 2.8%; Score 8; DB 1; Length 317;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 137 RARRAIA 144  
|||||  
DB 160 RARRAIA 167

RESULT 9  
FDBH\_METJA STANDARD; PRT: 379 AA.  
ID FDBH\_METJA  
AC Q60316;

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase beta chain (EC 1.2.1.2).
GN M00005.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischman R.D.,
RA Bult C.J., Blake O., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Peterson J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhman J.L., Nguyen D.,
RA Ullrich T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Frazer C.M., Smith H.O., Moese C.R., Venter J.C.;
RA *Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RT Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.
CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (PROBABLE).
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -1- SIMILARITY: STRONG, TO THE BETA SUBUNIT OF M.THERMOTUTROPHICUM
CC FDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67459; AAB97986.1; -
DR TIGR: M00005; -
DR InterPro: IPR001450; 4Fe4S-ferredoxin.
DR Pfam: PF00037; Fer4; 1.
DR PROSITE: PS00198; 4Fe4S-FERREDOXIN; 2.
KW Hypothetical protein; Oxidoreductase; NAD; Electron transport;
KW Iron-sulfur; 4Fe-4S; Complete proteome.
KM METAL 280 280 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 283 283 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 286 286 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 290 290 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 330 330 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 333 333 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 336 336 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 340 340 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 379 AA; 43014 MM; 9C257CCAD5547F5A CRC64;

Query Match 2.8%; Score 8; DB 1; Length 379;
Best local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 217 LVNGVIAL 224
DB 35 LVNGVIAL 42

```

```

RESULT 10
ZP3_MESAU STANDARD; PRT; 422 AA.
AC P23491;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

```

```

DE Zona pellucida sperm-binding protein 3 precursor (Zona pellucida
DE glycoprotein ZP3) (Sperm receptor) (Zona pellucida protein C).
GN ZP3.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ovary;
RX MEDLINE=91078540; PubMed=2257975;
RA Kinloch R.A., Ruiz-Seller B., Wasserman P.M.;
RT "Genomic organization and polypeptide primary structure of zona
RT pellucida glycoprotein hZP3, the hamster sperm receptor.";
RL Dev. Biol. 142:414-421(1990).
CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC SPECIES-SPECIFICITY OF THE INSEMINATION.
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
CC matrix.
CC -----
CC -1- TISSUE SPECIFICITY: OOCYTES.
CC -1- DEVELOPMENTAL STAGE: GROWING OOCYTES.
CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M63629; AAA37079.1; -
DR InterPro: IPR001507; Endoglin/CD105.
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPELLUCIDA.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix.
KM SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 422 ZONA PELLUCIDA SPERM-BINDING PROTEIN 3.
FT DOMAIN 23 386 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 387 407 POTENTIAL.
FT DOMAIN 408 422 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 306 ZP.
FT DOMAIN 119 158 PRO-RICH.
FT DOMAIN 208 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 422 AA; 45827 MM; D0F95BE7F887E01 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 422;
Best local Similarity 100.0%; Pred. No. 8.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 59 LAIGGLA 66
DB 386 LAIGGLA 393

```

```

RESULT 11
CYDC_HAEIN STANDARD; PRT; 576 AA.
AC P45081;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Transport ATP-binding protein cydc.
GN CYDC OR H1156.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kesteven K., Sutton G., Fitch W., Fields C.A., Merrick J.M.,
RA McEwen K., Sutton G., Fitch W., Fields C.A., Merrick J.M.,
RA Scott J.D., Shiley R., Liu L.-I., Glodok A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD Science 269:496-512(1995).
CC -I- FUNCTION: SOMEHOW INVOLVED IN THE CYTOCHROME D BRANCH OF AEROBIC
CC RESPIRATION. SEEMS TO BE A COMPONENT OF A TRANSPORT SYSTEM
CC (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32795; AAC22811.1; -.
DR TIGR: H1156; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001140; ABCtransportr.
DR Pfam: PF00005; ABC_tran; 1.
DR Pfam: PF00664; ABC_membrane; 1.
DR PRODOM: PD000006; ABC_transportr; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport; Transmembrane; Inner membrane;
KW Complete proteome.
KW TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 155 175 POTENTIAL.
FT TRANSMEM 244 264 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT NP_BIND 372 379 ATP (POTENTIAL).
SQ SEQUENCE 576 AA; 64831 MM; A9ACDB8B9B294A1B3 CRC64;

```

```

Query Match 2.8%; Score 8; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 58 PLALGIGL 65
Db 159 PLALGIGL 166

```

```

RESULT 12
GGT5_HUMAN STANDARD; PRT; 586 AA.
AC P36269; 096PCL;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-glutamyltranspeptidase 5 precursor (EC 2.3.2.2) (Gamma-
GN glutamyltransferase 5) (GGT-Rel).
GN GGT1A1 OR GGT5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91296809; PubMed=1676842;
RA Heisterkamp N., Rajpert-De Meyts E., Uribe L., Forman H.J.,
RA Groffen J.;
RT "Identification of a human gamma-glutamyl cleaving enzyme related to,
RT but distinct from, gamma-glutamyl transpeptidase."
RT Proc. Natl. Acad. Sci. U.S.A. 88:6303-6307(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RX Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: CAN HYDROLYZE THE GAMMA-GLUTAMYL MOIETY OF GLUTATHIONE;
CC AS WELL AS CONVERT LEUKOTRIENE C4 TO LEUKOTRIENE D4.
CC -I- CATALYTIC ACTIVITY: (5-L-glutamyl)-peptide + an amino acid -
CC peptide + 5-L-glutamyl-amino acid.
CC -I- SUBUNIT: THIS ENZYME CONSISTS OF TWO POLYPEPTIDE CHAINS, WHICH ARE
CC SYNTHESIZED IN PRECURSOR FORM FROM A SINGLE POLYPEPTIDE (BY
CC SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -I- SIMILARITY: STRONG, TO OTHER GGT ENZYMES AND TO GL-7ACA ACYLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64099; AA58503.1; -.
DR EMBL: BC011362; AAH11362.1; -.
DR PIR: A41125; A41125.
DR MEROPS: T03.002; -.
DR Gene: HGNC:4260; GGT1A1.
DR MIM: 137168; -.
DR InterPro: IPR00101; Gglutnsptdase.
DR Pfam: PF01019; G-glu-transpept; 1.
DR PRINTS: PR01210; GGTTRANSPASE.
DR PROSITE: PS00462; G-GLU-TRANSPEPTIDASE; 1.
DR Transferrase; Acyltransferase; Zymogen; Glycoprotein; Transmembrane;
KW Glutathione biosynthesis; Signal-anchor.
KW CHAIN 1 387 GAMMA-GLUTAMYLTRANSPPEPTIDASE 5 HEAVY
FT CHAIN 388 586 GAMMA-GLUTAMYLTRANSPPEPTIDASE 5 LIGHT
FT FT CHAIN (POTENTIAL).
FT FT GAMMA-GLUTAMYLTRANSPPEPTIDASE 5 LIGHT
FT FT CHAIN (POTENTIAL).
FT FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (PROBABLE).
FT FT LOMENAL (POTENTIAL).
FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CONFLICT 330 330 R -> K (IN REF. 2).
FT FT CONFLICT 437 437 W -> R (IN REF. 2).
SQ SEQUENCE 586 AA; 62319 MM; 1B8543CB0934B16B CRC64;

```

```

Query Match 2.8%; Score 8; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 61 LGIGLALA 68  
 DB 14 LGIGLALA 21

## RESULT 13

HSP1\_SAGIM STANDARD; PRT; 49 AA.  
 AC P24714;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sperm protamine P1 (Cysteine-rich protamine).  
 GN PRM1.  
 OS Saguinus imperator (Tamarin).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.  
 OX NCBI\_TaxID=9491;

## SEQUENCE FROM N.A.

RT TISSUE=Liver;  
 RX MEDLINE=92051332; PubMed=1840669;  
 RA Queralt R., Oliva R.;

RT "Protamine 1 gene sequence from the primate Saguinus imperator  
 isolated with PCR using consensus oligonucleotides.";  
 RL Nucleic Acids Res. 19:5786-5786(1991).

CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC Sperm DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE  
 CC DNA-HELIX (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: TESTIS.

CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: X61678; CAA3853.1; -

DR PIR: S22582; S22582.

DR InterPro: IPR000221; Protamine\_P1.

DR Pfam: PF00260; protamine\_P1.1.

DR PROSITE: PS00048; PROTAMINE\_P1.1.

KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW Testis; DNA condensation; Nuclear protein.

FT INIT\_MET 0  
 FT SEQUENCE 49 AA; 6545 MW; 8389C403F5B207F6 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 ORRRGR 48  
 DB 17 ORRRGR 23

## RESULT 14

HSP1\_DIDMA STANDARD; PRT; 57 AA.  
 AC P35305;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sperm protamine P1.  
 GN PRM1.

OS Didelphis marsupialis virginiana (North American opossum), and  
 OS Monodelphis domestica (Short-tailed grey opossum).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
 OX NCBI\_TaxID=9267, 13616;

## SEQUENCE FROM N.A.

RT SPECIES=D.marsupialis;  
 RX MEDLINE=93345500; PubMed=8344286;  
 RA Winkfein R.J., Nishikawa S., Connor W., Dixon G.H.;  
 RT "Characterization of a marsupial sperm protamine gene and its  
 RT transcripts from the North American opossum (Didelphis  
 RT marsupialis).";  
 RL Eur. J. Biochem. 215:63-72(1993).

## SEQUENCE FROM N.A.

RT SPECIES=M.domestica;  
 RX MEDLINE=95215351; PubMed=7700877;  
 RA Retief J.D., Kratjewski C., Westerman M., Winkfein R.J., Dixon G.H.;  
 RT "Molecular phylogeny and evolution of marsupial protamine P1 genes.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 259:7-14(1995).

CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC Sperm DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: TESTIS.

CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: L17007; AAA02812.1; -

DR EMBL: X74004; CAA52193.1; -

DR EMBL: U55448; AAA74612.1; -

DR PIR: S34045; S34045.

DR InterPro: IPR000221; Protamine\_P1.

DR Pfam: PF00260; protamine\_P1.1.

DR PROSITE: PS00048; PROTAMINE\_P1.1.

KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW Testis; DNA condensation; Nuclear protein.

FT INIT\_MET 0  
 FT BY SIMILARITY 0

FT SEQUENCE 57 AA; 7810 MW; 283715B280214E52 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 DB 34 RRRGRG 40

## RESULT 15

A62F\_DROME STANDARD; PRT; 115 AA.  
 AC 046202;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Accessory gland protein Acp62F precursor.  
 GN ACP62F OR CG1262.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;

RT "Characterization of a marsupial sperm protamine gene and its  
 RT transcripts from the North American opossum (Didelphis  
 RT marsupialis).";  
 RL Eur. J. Biochem. 215:63-72(1993).

RT "Characterization of a marsupial sperm protamine gene and its  
 RT transcripts from the North American opossum (Didelphis  
 RT marsupialis).";  
 RL Eur. J. Biochem. 215:63-72(1993).

RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,  
 RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tram U.K.;  
 RT "New genes for male accessory gland proteins in *Drosophila*  
 RT melanogaster.";  
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
 RA Balilew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rainett K., Remington K., Saunders R.D., Scheefel F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu G., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [13]  
 RP SEQUENCE OF 7-111 FROM N.A.  
 RC STRAIN-Various strains;  
 RX MEDLINE=20556153; PubMed=11102381;  
 RA Begun D.J., Whitley P., Todd B.L., Waldrup-Dall H.M., Clark A.G.;  
 RT "Molecular population genetics of male accessory gland proteins in  
 RT *Drosophila*.";  
 RL Genetics 156:1879-1888(2000).  
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN  
 CC MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL  
 CC FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO  
 CC AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STORAGE  
 CC AND EGG RELEASE.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: SEMINAL FLUID.  
 CC -1- SIMILARITY: SOME, TO P. NIGRIVENTER TX2-6.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; 085763; AAB96387.1; -.

DR EMBL; AE003475; AAF47683.1; -  
 DR EMBL; AY010608; AAG35367.1; -  
 DR EMBL; AY010609; AAG35368.1; -  
 DR EMBL; AY010610; AAG35369.1; -  
 DR EMBL; AY010611; AAG35370.1; -  
 DR EMBL; AY010612; AAG35371.1; -  
 DR EMBL; AY010613; AAG35372.1; -  
 DR EMBL; AY010614; AAG35373.1; -  
 DR EMBL; AY010615; AAG35374.1; -  
 DR EMBL; AY010616; AAG35375.1; -  
 DR EMBL; AY010617; AAG35376.1; -  
 DR FLYBASE; FBgn0020509; ACP62F.  
 DR InterPro; IPR002919; TIL\_Cysrich.  
 DR Pfam; PF01826; TIL\_1.  
 KW Behavior; signal.  
 FT SIGNAL 1 24  
 FT CHAIN 25 115  
 FT SEQUENCE 115 AA; 12570 MW; 4326A6F6C32291D CRC64;  
 SQ  
 Query Match 2.5%; Score 7; DB 1; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Gaps 0;  
 QY 68 ACIGLL 74  
 Db 11 ACIGLL 17  
 |||||  
 RESULT 16  
 Y151\_ECOLI  
 ID Y151\_ECOLI STANDARD; PRT; 118 AA.  
 AC P03838;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Insertion element IS5 very hypothetical 12 kda protein.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxId=562;  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82028652; PubMed=6269959;  
 RA Engler J.A., van Bree M.P.;  
 RT "The nucleotide sequence of IS5 from *Escherichia coli*.";  
 RL Gene 14:165-174(1981).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82028652; PubMed=6269958;  
 RA Engler J.A., van Bree M.P.;  
 RT "The nucleotide sequence and protein-coding capability of the  
 RT transposable element IS5.";  
 RL Gene 14:155-163(1981).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,  
 RA Davis K., Federpiet N., Hyman R., Kalman S., Komp C., Kurdi O.,  
 RA Ishikari D., Lew H., Lin D., Namath A., Oetner P., Roberts D.,  
 RA Davis R.W.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,  
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R.,  
 RA Ichihara S., Mizuno T., Makino K., Nakata A., Yura T., Sempel G.,  
 RA Mizobuchi K.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=97251357; PubMed=9097039;  
 RA Alba H., Bada T., Fujita K., Hayashi K., Inada T., Isono K.,

```

RA Itch T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizouchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampel G., Seki Y., Sivasundaram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horieuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01734; -; NOT_ANNOTATED_CDS.
DR EMBL: U70214; AAB086B0.1; -.
DR EMBL: D83536; -; NOT_ANNOTATED_CDS.
DR EMBL: D90771; BAA14925.1; -.
DR EMBL: D90772; BAA14935.1; -.
DR EMBL: D90831; BAA15715.1; -.
DR EMBL: D90841; BAA15872.1; -.
DR EMBL: D90847; BAA15958.1; -.
DR EMBL: D90848; BAA15963.1; -.
DR PIR: A04466; IECSB.
KW Hypothetical protein; Transposable element.
SQ SEQUENCE 118 AA; 12270 MW; 348014FAC765058E CRC64;

Query Match 2.5%; Score 7; DB 1; Length 118;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 DCGAVRQ 29
DB 107 DCGAVRQ 113

RESULT 17
IL13_MOUSE STANDARD; PRT; 131 AA.
ID IL13_MOUSE
AC P20109;
DT 01-FEB-1991 (Rel. 17; Created)
DT 01-FEB-1991 (Rel. 17; Last sequence update)
DT 15-JUL-1998 (Rel. 36; Last annotation update)
DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).
GN IL13 OR IL-13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89093958; PubMed=2521353;
RA Brown K.D., Zurawski S.M., Mosmann T.R., Zurawski G.;
RT "A family of small inducible proteins secreted by leukocytes are
RT members of a new superfamily that includes leukocyte and
RT fibroblast-derived inflammatory agents, growth factors, and
RT indicators of various activation processes.";
RL J. Immunol. 142:679-687(1989).
CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.
CC SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
CC MAY BE CRITICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONSES
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23504; AAA40149.1; -.
DR PIR: E30552; E30552.
DR HSP: P35225; 3ITR.
DR MGD: MGI:96541; 1113.
DR InterPro: IPR003634; Interleukin_13.
DR InterPro: IPR001325; Interleukin_4_13.
DR Pfam: PF03487; Interleukin_13; 1.
DR ProDom: PD015987; Interleukin_13; 1.
DR SMART: SM00190; IL4_13; 1.
DR PROSITE: PS00838; INTERLEUKIN_4_13; 1.
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 1 131 INTERLEUKIN_13.
FT DISULFID 51 79 BY SIMILARITY.
FT DISULFID 67 93 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 131 AA; 14107 MW; 954F93F105713FED CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 71
DB 9 LALACLG 15

RESULT 18
IL13_RAT STANDARD; PRT; 131 AA.
ID IL13_RAT
AC P42203;
DT 01-NOV-1995 (Rel. 32; Created)
DT 01-NOV-1995 (Rel. 32; Last sequence update)
DT 15-JUL-1998 (Rel. 36; Last annotation update)
DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).
GN IL13 OR IL-13.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Kidney cortex;
RX MEDLINE=94092138; PubMed=7916615;
RA Laktis F.G., Cruet E.N.;
RT "Cloning of rat Interleukin-13 (IL-13) cDNA and analysis of IL-13
RT gene expression in experimental glomerulonephritis.";
RL Biochem. Biophys. Res. Commun. 197:612-618(1993).
CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.
CC SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
CC MAY BE CRITICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONSES
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U26913; AAA16478.1; -.
DR HSP: P35225; 3ITR.
DR InterPro: IPR003634; Interleukin_13.
DR InterPro: IPR001325; Interleukin_4_13.

```

```

DR Pfam: PF03487; Interleukin_13; 1.
DR Prodom: PD015987; Interleukin_13; 1.
DR SMART: SM00190; IL4_13; 1.
DR PROSITE: PS00838; INTERLEUKIN_4_13; 1.
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 131 INTERLEUKIN_13.
FT DISULFID 52 80 BY SIMILARITY.
FT DISULFID 68 94 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 131 AA; 14093 MW; E508CAB8DE8C201 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LALACLG 71
Db 9 LALACLG 15

RESULT 19
YK01_PYRHO STANDARD; PRT; 147 AA.
AC 057781;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein PH2001.
GN PH2001.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RA MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamaya M., Ohnuki Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki Y., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuza H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000001; BAA31940.1; -
DR EMBL: AP000007; BAA31940.1; JOINED.
DR EMBL: AP000007; BAA31943.1; -
DR EMBL: AP000001; BAA31943.1; JOINED.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
SQ SEQUENCE 147 AA; 15324 MW; 247ED12FCECF265B9 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 LACIGLL 73

```

```

Db 45 LACIGLL 51

RESULT 20
TNFC_PIG STANDARD; PRT; 150 AA.
AC 09TSV8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
DE necrosis factor ligand superfamily member 3) (Fragment).
GN LTB OR TNFSF3 OR TNFC.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=large white; Tissue=Fibroblast;
RX MEDLINE=21108615; PubMed=11169259;
RA Chardon P., Rogel-Galliard C., Cattolico L., Duprat S., Vaiman M.,
RA Renard C.;
RT "Sequence of the swine major histocompatibility complex region
RT containing all non-classical class I genes.";
RL Tissue Antigens 57:55-65(2001).
CC CC
CC -I- FUNCTION: Cytokine that binds to LTB/ TNFSF3. May play a specific
CC role in immune response regulation. Provides the membrane anchor
CC for the attachment of the heterotrimeric complex to the cell
CC surface.
CC -I- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or
CC (less prevalent) two LTA and one LTB subunits.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein (by similarity).
CC -I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ251914; CAB63851.1; -
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_Family.
DR Pfam: PF00229; TNF; 1.
DR PRINTS: PR01234; TNFCROSISFCT.
DR Prodom: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS00499; TNF_2; 1.
KW Cytokine; Cytotoxin; Glycoprotein.
FT NON-TER 1 1
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16423 MW; F5C4C657658B48 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GLYLYYC 200
Db 37 GLYLYYC 43

RESULT 21
RA05_ORYSA STANDARD; PRT; 157 AA.
ID RA05_ORYSA
AC Q01881;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

DT 01-JUN-1994 (rel. 29, last annotation update)
DE seed allergenic protein RA5 precursor.
GN RA5.
OS Oryza sativa (Rice).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae:
OC Ehrhartoideae: Oryzaceae: Oryza.
OX NCBI_TaxID=4530;
RN RP
RP SEQUENCE FROM N.A.
RC TISSUE=Seed:
RC MEDLINE=93144699; PubMed=7678765;
RA Adachi T., Izumi H., Yamada T., Tanaka K., Takeuchi S.,
RA Nakamura R., Matsuda T.;
RT "Gene structure and expression of rice seed allergenic proteins
RT belonging to the alpha-amylase/trypsin inhibitor family.";
RL Plant Mol. Biol. 21:239-248(1993).
CC -1- FIVE DISULFIDE BONDS ARE PRESENT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPsin/ALPHA-AMYLASE INHIBITOR
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D11430; BAA01996.1; -
DR PIR: S31078; S31078.
DR HSSP: P01085; IHSS.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; TRY/amy1_inhbr.
DR Pfam: PF00234; tryp_alpha_aml1; 1.
DR PRINTS: PRO00808; AMLASEINHBR.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW Allergen; Multigene family; Signal.
FT SIGNAL 1 26
FT CHAIN 27 157
FT SEQUENCE 157 AA; 17118 MW; C8A5495FBFB399E6 CRC64;
SQ
Query Match 2.5%; Score 7; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 72 LLLAVVS 78
DB 11 LLLAVVS 17

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
RT Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O.,
RA Lashari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7.";
RT Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka H., Tooe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RP SEQUENCE OF 72-161 FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=91100302; PubMed=1987124;
RA Dicker I.B., Seetharam S.R.;
RT "Cloning and nucleotide sequence of the fira gene and the fira200(Ts)
RT allele from Escherichia coli.";
RL J. Bacteriol. 173:334-344(1991).
RN [8]
RP SEQUENCE OF 21-32.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [9]
RP SIMILARITY TO S.TYPHIMURIUM OMPH.
RX MEDLINE=90201355; PubMed=2318304;
RA Hirvas L., Coleman J., Koski P., Vaara M.;
RT "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane
RT constituent?";
RL FEBS Lett. 262:123-126(1990).
CC -1- SUBUNIT: HOMOTETRAMER.

```



```

CC -1- SUBCELLULAR LOCATION: EITHER IN THE NUCLEOID (CHROMATIN) OR IN
CC THE OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE OMPH/HMPA FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M21118; AAA24630.1; -
DR EMBL: AE000127; AAC73289.1; -
DR EMBL: D83536; BAA77853.1; -
DR EMBL: U70214; AAB08607.1; -
DR EMBL: AE005194; AAG54480.1; -
DR EMBL: AP002550; BAB33603.1; -
DR EMBL: X54797; CAA38567.1; -
DR EMBL: X75465; CAA53207.1; -
DR PIR: JTO304; DNEC17.
DR PIR: S13728; S13728.
DR Ecogene; BG10455; hmpa.
DR DNA-binding; Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 20
FT CHAIN 21 161 HISTONE-LIKE PROTEIN HLP-1.
SQ SEQUENCE 161 AA; 17688 MW; 2A966BBD83FE675 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GLGLALA 68
Db 9 GLGLALA 15

RESULT 23
YLMA CAEL
ID YLMA CAEL STANDARD; PRT: 170 AA.
AC P34378;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein D2007.4 in chromosome III.
GN D2007.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Pezodermidae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Lalister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showken R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J.,
RA Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: L16560; AAA27999.1; -
DR PIR: S44789; S44789.
DR Wormpep: D2007.4; CE00129.
KW Hypothetical protein.
SQ SEQUENCE 170 AA; 19396 MW; 22301D7C65638135 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 VLALRCL 227
Db 94 VLALRCL 100

RESULT 24
LACB_BUBBU
ID LACB_BUBBU STANDARD; PRT: 180 AA.
AC P02735; O62822;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-lactoglobulin precursor (Beta-LG).
GN LGB.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_Taxid=89462;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE-Mammary gland;
RX MEDLINE=99304500; PubMed=10376212;
RA Das P., Jain S., Nayak S., Apparo K.B.C., Torey S.M., Garg L.C.;
RT "Molecular cloning and sequence analysis of the cDNA encoding
RT beta-lactoglobulin in Bubalus bubalis.";
RL DNA Seq. 10:105-108(1999).
RN 12
RP SEQUENCE OF 19-180.
RA Kolde H.-J., Liberatori J., Brauntzer G.;
RT "The amino acid sequence of the water buffalo beta-lactoglobulin.";
RL Milchwissenschaft 36:83-86(1981).
CC -1- FUNCTION: Primary component of whey, it binds retinol and is
CC probably involved in the transport of that molecule.
CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS
CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Synthesized in mammary gland and secreted
CC in milk.
CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.
CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AJ005429; CAA06532.1; -
DR PIR: A03219; IGBUI.
DR HSSP: P02754; IBSQ.
DR InterPro: IPR002345; Lipocalin.
DR InterPro: IPR000566; Lipocalin_cytrFABP.
DR Pfam: PF00061; Lipocalin; 1.
DR PRINTS: PR00179; LIPOCALIN.
DR PROSITE: PS00213; LIPOCALIN; 1.
KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.

```

FT SIGNAL 1 18  
FT CHAIN 19 180  
FT DISULFID 84 178  
FT DISULFID 124 137  
FT DISULFID 124 139  
SQ SEQUENCE 180 AA; 20023 MW; 6836C97B2C2E33CF CRC64;  
ALTERNATE:  
Query Match 2.5%; Score 7; DB 1; Length 180;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 63 LGALAC 69  
DB 8 LGALAC 14  
RESULT 25  
LACB\_CAPHT STANDARD; PRT; 180 AA.  
ID LACB\_CAPHT  
AC P02756;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Beta-lactoglobulin precursor (beta-LG).  
GN LGB.  
OS Capra hircus (Goat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Capra.  
OX NCBI\_TaxID=9925;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sp. aegagrus; TISSUE=Mammary gland;  
RX MEDLINE=94042559; PubMed=8226387;  
RA Folch J.M., Coll A., Sanchez A.;  
RT "Cloning and sequencing of the cDNA encoding goat beta-lactoglobulin.";  
RT J. Anim. Sci. 71:2832-2832(1993).  
RN (2)  
RP SEQUENCE FROM N.A.  
RA Kim J., Kim A., Kim J., Yu M.;  
RT Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
RN (3)  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95213451; PubMed=7699130;  
RA Folch J., Coll A., Sanchez A.;  
RT "Complete sequence of the caprine beta-lactoglobulin gene.";  
RT J. Dairy Sci. 77:3493-3497(1994).  
RN (4)  
RP SEQUENCE OF 19-180.  
RX MEDLINE=80070611; PubMed=511095;  
RA Preaux G., Brauntz G., Schrank B., Stangl A.;  
RT "The amino acid sequence of goat beta-lactoglobulin.";  
RT Hoppe-Seyler's Z. Physiol. Chem. 360:1595-1604(1979).  
RN (5)  
RP FUNCTION: Primary component of whey, it binds retinol and is probably involved in the transport of that molecule.  
RN (6)  
RP SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.  
RN (7)  
RP TISSUE SPECIFICITY: Secreted.  
RN (8)  
RP TISSUE SPECIFICITY: Synthesized in mammary gland and secreted in milk.  
RN (9)  
RP MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.  
RN (10)  
RP SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; X58471; CAA41385.1; -  
DR EMBL; Z19569; CAA79623.1; -  
DR EMBL; Z19570; CAA79623.1; -  
DR EMBL; Z33881; CAA83946.1; -  
DR PIR; A03220; LGCT  
DR PIR; S14507; S14507.  
DR PIR; S14507; S14507.  
DR HSSP; P02754; 1B00.  
DR HSSP; P02754; 1B00.  
DR InterPro; IPR002345; Lipocalin.  
DR InterPro; IPR000566; Lipocalin.  
DR Pfam; PF00061; Lipocalin; 1.  
DR PRINTS; PR00179; LIPOCALIN.  
DR PROSITE; PS00213; LIPOCALIN; 1.  
KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.  
FT SIGNAL 1 18  
FT CHAIN 19 180  
FT DISULFID 84 178  
FT DISULFID 124 137  
FT DISULFID 124 139  
SQ SEQUENCE 180 AA; 19975 MW; C2449BB02A1A80F1 CRC64;  
ALTERNATE:  
Query Match 2.5%; Score 7; DB 1; Length 180;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 63 LGALAC 69  
DB 8 LGALAC 14  
RESULT 26  
LACB\_SHEEP STANDARD; PRT; 180 AA.  
ID LACB\_SHEEP  
AC P02757;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Beta-lactoglobulin 1/B, 2/A, and 3/C precursor.  
OS Ovis aries (Sheep), and  
OS Ovis orientalis musimon (Mouflon).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940, 9938;  
RN (1)  
RP SEQUENCE FROM N.A. (BLG 1 AND 2).  
RC SPECIES=Sheep;  
RX MEDLINE=88172489; PubMed=3351935;  
RA Ali S., Clark A.J.;  
RT "Characterization of the gene encoding ovine beta-lactoglobulin. RT similarity to the genes for retinol binding protein and other secretory proteins.";  
RT J. Mol. Biol. 199:415-426(1988).  
RN (2)  
RP SEQUENCE FROM N.A. (BLG 1).  
RC SPECIES=Sheep;  
RX MEDLINE=87049827; PubMed=3096387;  
RA Gave P., Hue-Delahaie D., Mercier J.-C., Soulier S., Vilotte J.-L., Furet J.-P.;  
RT "Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland.";  
RT Biochimie 68:1097-1107(1986).  
RN (3)  
RP SEQUENCE FROM N.A. (BLG 1).  
RC SPECIES=Sheep;  
RX MEDLINE=89057492; PubMed=3194215;  
RA Harris S., Ali S., Anderson S., Archibald A.L., Clark A.J.;  
RT "Complete nucleotide sequence of the genomic ovine beta-lactoglobulin gene.";  
RT Nucleic Acids Res. 16:10379-10380(1988).  
RN (4)  
RP SEQUENCE FROM N.A. (BLG 1 AND 2).  
RC SPECIES=Sheep;  
-----

RX MEDLINE=91007276; PubMed=1976573;  
 RA Ali S., McClenaghan M., Simons J.P., Clark A.J.;  
 RT "Characterisation of the alleles encoding ovine beta-lactoglobulins A  
 RT and B.";  
 RL Gene 91:201-207(1990).  
 RN [5]  
 RP SEQUENCE OF 19-180 (BLG 2).  
 RC SPECIES=Sheep;  
 RX MEDLINE=80219294; PubMed=6155855;  
 RA Preaux G., Braunitzer G., Kolde H.-J.;  
 RT "Primary structure of ovine beta-lactoglobulin.";  
 RL Arch. Int. Physiol. Biochim. 88:B45-B46(1980).  
 RN [6]  
 RP SEQUENCE OF 19-180 (BLG 3).  
 RC SPECIES=Sheep;  
 RX MEDLINE=89374823; PubMed=2775495;  
 RA Ehardt G., Godovac-Zimmermann J., Conti A.;  
 RT "Isolation and complete primary sequence of a new ovine wild-type  
 RT beta-lactoglobulin C.";  
 RL Biol. Chem. Hoppe-Seyler 370:757-762(1989).  
 RN [7]  
 RP SEQUENCE OF 19-180 (BLG B).  
 RC SPECIES=O. musimon;  
 RX MEDLINE=88106996; PubMed=3426802;  
 RA Godovac-Zimmermann J., Conti A., Napolitano L.;  
 RT "The complete amino-acid sequence of dimeric beta-lactoglobulin from  
 RT mouflon (Ovis ammon musimon) milk.";  
 RL Biol. Chem. Hoppe-Seyler 368:1313-1319(1987).  
 CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT  
 CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF  
 CC THAT MOLECULE.  
 CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS  
 CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.  
 CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.  
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04520: CAA28204.1: -;  
 DR EMBL: X12817: CAA31305.1: -;  
 DR EMBL: X07004: CAA30059.1: ALT\_SEQ.  
 DR EMBL: X07005: CAA30059.1: JOINED.  
 DR EMBL: X07006: CAA30059.1: JOINED.  
 DR EMBL: X07007: CAA30059.1: JOINED.  
 DR EMBL: X07008: CAA30059.1: JOINED.  
 DR EMBL: X07009: CAA30059.1: JOINED.  
 DR EMBL: M32236: AAA31510.1: -;  
 DR EMBL: M32233: AAA31510.1: JOINED.  
 DR EMBL: M32232: AAA31510.1: JOINED.  
 DR EMBL: M32235: AAA31510.1: JOINED.  
 DR PIR: A03221: LGSF.  
 DR PIR: S00132: S00132.  
 DR PIR: A30011: A30011.  
 DR PIR: B30011: B30011.  
 DR PIR: A25136: A25136.  
 DR PIR: J00748: J00748.  
 DR PIR: J00749: J00749.  
 DR PIR: S02136: S02136.  
 DR PIR: S04955: S04955.  
 DR HSSP: P02754: IBSQ.  
 DR InterPro: IPR002345: Lipocalin.  
 DR Pfam: PF00061: Lipocalin\_cyFABP.  
 DR PRINTS: PR00179: LIPOCALIN.  
 DR PROSITE: PS00213: LIPOCALIN.  
 DR Milk: Whey; Retinol-binding; Transport; Signal; Lipocalin.  
 FT SIGNAL 1 18

FT CHAIN 19 180 BETA-LACTOGLOBULIN.  
 FT DISULFID 84 178  
 FT DISULFID 124 137  
 FT DISULFID 124 139  
 FT VARIANT 38 38  
 FT VARIANT 166 166  
 SQ SEQUENCE 180 AA; 19921 MW; BABCB2E89E757333 CRC64;  
 Query Match 2.5%; Score 7; DB 1; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 63 LGALAC 69  
 Db 8 LGALAC 14  
 RESULT 27  
 ID RNFE\_SALTY STANDARD; PRT; 230 AA.  
 AC ORKEX9;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Electron transport complex protein rnfe.  
 GN RNFE OR STM1454 OR STY1668.  
 OS Salmonella typhimurium, and  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_Taxid=602, 601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Salmonella; STRAIN=LT2 / SGC1412 / ATCC 700720;  
 RX MEDLINE=21534947; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Flora L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RT LT2.";  
 RL Nature 413:852-856(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Salmonella; STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
 RA Baker S., Basham D., Brooks R., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies K.M., Dowd L., White N., Farrar J.,  
 RA Felzwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagers K.,  
 RA Krogli M., Larsen T.S., Leather S., Moule S., O'Garra P., Parry C.,  
 RA Whitehead S., Barrall B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RT enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 CC -1- FUNCTION: May be part of a membrane complex involved in electron  
 CC transport (By similarity).  
 CC -1- SUBUNIT: Composed of at least six subunits; rnfa, rnfb, rnfc,  
 CC rnfd, rnfe and rnfg (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

CC -----
DR EMBL: AE008763; AAL20376.1; -.
DR EMBL: AL627271; CAD01913.1; -.
DR StycGene: SG272727; rnfE.
DR InterPro: IPR003667; Rnf_Nqr.
DR Pfam: PF02508; Rnf-Nqr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 34 56 POTENTIAL.
FT TRANSMEM 69 87 POTENTIAL.
FT TRANSMEM 91 113 POTENTIAL.
FT TRANSMEM 126 148 POTENTIAL.
FT TRANSMEM 183 205 POTENTIAL.
SQ SEQUENCE 230 AA: 24318 MW: E198B4CEA13F249E CRC64;

Query Match 2.5%; Score 7; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLA 66
DB 38 ALGLGLA 44

RESULT 28
RNF_ECO57 STANDARD: PRT: 231 AA.
ID RNF_ECO57
AC P58344;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF OR Z2642 OR ECS2341.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RA MEDLINE-21074935; PubMed-11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Grodbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,
RA Apodaca J., Aanantaraman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RA Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RA MEDLINE-21156231; PubMed-11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
RA Ikeda T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuwara S., Shiba T., Hattori M., Shingawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC -----
CC or send an email to license@isb-sib.ch).
DR EMBL: AE005386; AAG56621.1; -.
DR EMBL: AP002558; BAB35764.1; -.
DR InterPro: IPR003667; Rnf_Nqr.
DR Pfam: PF02508; Rnf-Nqr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 1 38 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 62 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 84 85 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 107 124 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 125 145 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 146 181 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 182 202 POTENTIAL.
FT TRANSMEM 203 231 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 231 AA: 24489 MW: DA42CA2D292604C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLA 66
DB 38 ALGLGLA 44

RESULT 29
RNF_ECOLI STANDARD: PRT: 231 AA.
ID RNF_ECOLI
AC P77179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF OR B1632.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / 551557; PubMed-9097039;
RA MEDLINE-97251357; PubMed-9097039;
RA Aliba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nasimoto H., Nishio Y.,
RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP TOPOLOGY.
RC STRAIN-K12 / JM109;
RA MEDLINE-99342054; PubMed-10411911;
RA Saef A., Johansson M., Wallin E., von Heijne G.;
RA "Divergent evolution of membrane protein topology: the Escherichia
RT coli RnfA and RnfE homologues.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 96:8540-8544(1999).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000258; AAC74704.1; -.
DR EMBL; D90806; BAA15386.1; -.
DR EMBL; D90807; BAA15393.1; -.
DR EMBL; D90808; BAA15416.1; -.
DR Ecogene; EG13938; rnfE.
DR InterPro; IPR003667; Rnf_Nqr.
DR Pfam; PF02508; Rnf_Nqr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT DOMAIN 1 38 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 PERIPLASMIC (POTENTIAL).
FT DOMAIN 60 62 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 63 83 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 84 85 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 86 106 PERIPLASMIC (POTENTIAL).
FT DOMAIN 107 124 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 125 145 PERIPLASMIC (POTENTIAL).
FT DOMAIN 146 181 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 182 202 PERIPLASMIC (POTENTIAL).
FT DOMAIN 203 231 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 231 AA; 24459 MW; CFA37A2D292604C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLGLA 66
| | | | |
DB 38 ALGLGLA 44

RESULT 30
RNC_COXBU STANDARD; PRT; 233 AA.
ID RNC_COXBU
AC P51837;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III).
GN RNC.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
CC Coxiella group; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CB9MIC7;
RX MEDLINE=95131751; PubMed=7830573;
RA Zuber M., Hoover T.A., Powell B.S., Court D.L.;
RT "Analysis of the rnc locus of Coxiella burnetii.";
RL Mol. Microbiol. 14:291-300(1994).
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.

```

```

CC -1- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L27436; AAA69690.1; -.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_3.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF00636; Ribonuclease_3; 1.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00353; RIBOC; 1.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; RNase_3_1; 1.
DR PROSITE; PS50142; RNase_3_2; 1.
KW Hydrolyase; Nuclease; Endonuclease; RNA-binding.
FT DOMAIN 4 126 RNASE III.
FT DOMAIN 204 220 DRBM.
SQ SEQUENCE 233 AA; 26229 MW; 1A1CB4AD96FD2FF CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 ARRLPLP 16
| | | | |
DB 164 ARRLPLP 170

RESULT 31
RNF_E_YERPE STANDARD; PRT; 233 AA.
ID RNF_E_YERPE
AC Q8ZED4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF_E OR YPO2240.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Tiltball R.W., Holden M.T.G.,
RA Prentice M.B., Sedalnia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AJ414151; CAC91046.1; -

DR InterPro: IPR003667; Rnf.Nqr.

DR Pfam: PF02508; Rnf.Nqr; 1.

KW Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 34 56 POTENTIAL.

FT TRANSMEM 69 87 POTENTIAL.

FT TRANSMEM 91 113 POTENTIAL.

FT TRANSMEM 126 148 POTENTIAL.

FT TRANSMEM 184 206 POTENTIAL.

SO SEQUENCE 233 AA; 24587 MW; 4918183358CB90 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLGLA 66  
Db 38 ALGLGLA 44  
|||||

RESULT 32

RNFE\_HAEIN STANDARD; PRT; 235 AA.

ID RNFE\_HAEIN 057020; P96346;

AC 057020; P96346;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Electron transport complex protein rnfE.

GN RNFE OR H11688.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI\_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Rd / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., McKelvey K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Ullrichback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;

RA "Whole-genome random sequencing and assembly of Haemophilus influenzae R.D.";

RT Science 269:496-512(1995).

RL -1- FUNCTION: May be part of a membrane complex involved in electron transport (By similarity).

CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC, rnfD, rnfE and rnfG (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).

CC -1- SIMILARITY: BELONGS TO THE NORDE/NRFAE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U32841; AAC23334.1; -

DR TIGR: H11688; -

DR InterPro: IPR003667; Rnf.Nqr.

DR Pfam: PF02508; Rnf.Nqr; 1.

KW Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 63 83 POTENTIAL.

FT TRANSMEM 93 113 POTENTIAL.

FT TRANSMEM 117 137 POTENTIAL.

FT TRANSMEM 152 172 POTENTIAL.

FT TRANSMEM 206 226 POTENTIAL.

SO SEQUENCE 235 AA; 25845 MW; C054FE96647837A CRC64;

Query Match 2.5%; Score 7; DB 1; Length 235;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLGLA 66  
Db 62 ALGLGLA 68  
|||||

RESULT 33

TN14\_MOUSE STANDARD; PRT; 239 AA.

ID TN14\_MOUSE 090YH9;

AC 090YH9;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 14.

GN TNFSF14 OR LIGHT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=10700230;

RA Tamada K., Shinozaki K., Chapoval A.I., Zhu G., Sica G., Flies D., Boone T., Hsu H., Fu Y.-X., Nagata S., Ni J., Chen L.;

RT "Modulation of T-cell-mediated immunity in tumor and graft-versus-host disease models through the LIGHT co-stimulatory pathway.";

RL Nat. Med. 6:283-289(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal liver;

RA Misawa K., Nosaka T., Kitamura T., Kojima T.;

RT "Murine LIGHT, a homologue of human LIGHT which is a member of TNF family.";

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Lymphoma;

RA Force W.R., Todd P.K., Mikayama T.;

RT "Mouse LIGHT: molecular genetics, ligand binding and expression.";

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Cytokine that binds to TNFRSF3/LTR. Binding to the decoy receptor TNFRSF6B modulates its effects. Activates NFkB and stimulates the proliferation of T cells.

CC -1- SUBUNIT: Homotrimer (By similarity).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By similarity).

CC -1- PTM: The soluble form derives from the membrane form by proteolytic processing.

CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AF123385; AAF76453.1; -

DR EMBL: AB029155; BAA88559.1; -

DR EMBL: AF227533; AAF36722.1; -

DR HSSP: P01375; 4TSV.

DR MGD: MGI:1355317; Tnfrsf14.

DR InterPro: IPR003636; TNF\_abc.  
DR InterPro: IPR000478; TNF\_family.  
DR Pfam: PF00229; TNF\_1.  
DR PRINTS: PR01234; TNECROSISFCT.  
DR ProDom: PD002012; TNF\_abc; 1.  
DR SMART: SM00207; TNF; 1.  
DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
DR PROSITE: PS0049; TNF\_2; 1.  
KM Cytochrome; Cytochrome; Transmembrane; Glycoprotein; Signal-anchor.  
FT CHAIN 1 239  
FT CHAIN 782 239  
FT CHAIN 1 37  
FT DOMAIN 1 37  
FT TRANSMEM 38 58  
FT TRANSMEM 59 239  
FT SITE 81 82  
FT DISULFID 152 187  
FT CARBOHYD 100 100  
FT CARBOHYD 191 191  
SQ SEQUENCE 239 AA; 26338 MW; 217874AC71AD6B3 CRC64;  
  
Query Match 2.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 121 NLRVPR 127  
Db 221 NLRVPR 227  
  
RESULT 34  
RNF\_ESEAE  
ID RNF\_ESEAE STANDARD; PRT; 240 AA.  
AC 09HYB5;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Electron transport complex protein rnfE.  
GN RNF\_E OR PA3494.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Hickley M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
Reizner J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen";  
RL Nature 406:959-964(2000).  
CC -1- FUNCTION: May be part of a membrane complex involved in electron  
CC transport (By similarity).  
CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,  
CC rnfD, rnfE and rnfG (By similarity).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL: AE004770; AAC06882.1;  
DR InterPro: IPR003667; Rnf\_Ngr.  
DR Pfam: PF02508; Rnf\_Ngr; 1.  
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 41 61  
FT TRANSMEM 71 91  
FT TRANSMEM 95 115  
FT TRANSMEM 130 150  
FT TRANSMEM 184 204  
SQ SEQUENCE 240 AA; 25794 MW; 3D90687ED462DBB2 CRC64;  
  
Query Match 2.5%; Score 7; DB 1; Length 240;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 60 ALGLGLA 66  
Db 40 ALGLGLA 46  
  
RESULT 35  
MOEB\_HAEIN  
ID MOEB\_HAEIN STANDARD; PRT; 243 AA.  
AC PA5211;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Molybdopterin biosynthesis protein moeb.  
GN MOEB OR CHLN OR H11449.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rd / KM20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
McKenny K., Sulten G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
Ufferback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
Rae L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
Venter J.C.;  
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd.";  
RL Science 269:496-512(1995).  
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR  
CC (MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE  
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING  
CC FACTOR (MOAF) (BY SIMILARITY).  
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.  
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Query Match 2.5%; Score 7; DB 1; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNROI 185  
 DB 10 LRYNROI 16

RESULT 36  
 TNFC\_HUMAN STANDARD; PRT: 244 AA.

AC 006643; P78370; Q99761;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor necrosis factor ligand superfamily member 3).  
 GN LTB OR TNFS3 OR TNFC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.  
 RC TISSUE=T-cell.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-Ek A., Lawton P., Demarini J., Tizard R., Chow E.P., Hession C., O'Brien-Greco B., Foley S.F., Ware C.F., "Lymphotoxin beta, a novel member of the TNF family that forms a heteromeric complex with lymphotoxin on the cell surface." Cell 72:847-856(1993).  
 RL [2]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=97445965; PubMed=9299492;  
 RA Matzocha K., Renard N., Charlot C., Bienvu J., Collier B., Salles G., "Identification of two lymphotoxin beta isoforms expressed in human lymphoid cell lines and non-Hodgkin's lymphomas." Biochem. Biophys. Res. Commun. 238:273-276(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Neville M.J., Miner C.M., Campbell R.D., "A new member of the immunoglobulin superfamily and a V-ATPase G subunit are amongst the predicted products of novel genes close to the TNF locus in the human MHC." Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A., Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J., Lasky S., Hood L., "Sequence of the human major histocompatibility complex class III region." Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Shina S., Tamiya G., Oka A., Inoko H., "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region." Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLU-70 AND PRO-111.  
 RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Toth E.J., Yi O., Nickerson D.A., Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to LTB/TNFSF3. May play a specific role in immune response regulation. Provides the membrane anchor for the attachment of the heterotrimeric complex to the cell surface. Isoform 2 is probably non-functional.  
 CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or (less prevalent) one LTB and two LTA subunits.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (potential).

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: SPLEEN AND THYMUS.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL; L11016; AAA99888.1; -;  
 DR EMBL; U08922; AAC51769.1; -;  
 DR EMBL; U79029; AAB37342.1; -;  
 DR EMBL; L11015; AAA36191.1; -;  
 DR EMBL; Y14768; CAA75069.1; -;  
 DR EMBL; AF129756; AAD18089.1; -;  
 DR EMBL; AP000505; BAB63395.1; -;  
 DR EMBL; AY070219; AAL49954.1; -;  
 DR EMBL; AY070219; AAL49955.1; -;  
 DR PIR; A46066; A46066.  
 DR HSSP; P01374; 1TNR.  
 DR GeneW; HGNC:6711; LTB.  
 DR MIM; 600978; -;  
 DR InterPro; IPR003636; TNF-abc.  
 DR InterPro; IPR000478; TNF-family.  
 DR Pfam; PF00229; TNF; 1.  
 DR PRINTS; PR01234; TNECROSISFCT.  
 DR ProDom; PD002012; TNF-abc; 1.  
 DR SMART; SM00207; TNF; 1.  
 DR PROSITE; PS00251; TNF\_1; 1.  
 DR PROSITE; PS0049; TNF\_2; 1.  
 KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;  
 KW Alternative splicing; Polymorphism.  
 FT DOMAIN 1 18  
 FT TRANSMEM 19 48  
 FT DOMAIN 1 18  
 FT TRANSMEM 19 48  
 FT DOMAIN 1 18  
 FT TRANSMEM 19 48  
 FT CARBOHD 222 222  
 FT VARSPLIC 53 77  
 FT VARSPLIC 78 244  
 FT VARIANT 70 70  
 FT VARIANT 111 111  
 FT CONFLICT 60 69  
 FT SEQUENCE 244 AA; 25390 MW; F41569459830ED4C CRC64;  
 SQ

Query Match 2.5%; Score 7; DB 1; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 194 GLYLYXC 200  
 DB 131 GLYLYXC 137

RESULT 37  
 MOEB\_ECOLI STANDARD; PRT: 249 AA.

AC P12282;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Molycoprotein biosynthesis protein moeb.  
 GN MOEB OR CHLN OR B0826.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.



```

OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88314906; PubMed=3045084;
RA Nohno T., Kasai Y., Saito T.;
RT Cloning and sequencing of the Escherichia coli chlE operon involved
RT in molybdopterin biosynthesis.;
RL J. Bacteriol. 170:4097-4102(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RA MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT The complete genome sequence of Escherichia coli K-12.;
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RA MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Mada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT A 7.8-Kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.;
RL DNA Res. 3:137-155(1996).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
CC (MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING
CC FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M21151; AAA23580.1; -
DR EMBL: AE000185; AAC73913.1; -
DR EMBL: D90720; BAA35514.1; -
DR EMBL: D90721; BAA35521.1; -
DR PIR: B32352; B32352.
DR EcoGene: EG10154; moeb.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; Thif_domain.
DR Pfam: PF00899; Thif. 1.
KW Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 249 AA; 26719 MW; 12C77082B3F39D7D CRC64;

Query Match 2.58; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNRQI 185
DB 10 LRYNRQI 16
|||||

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Molybdopterin biosynthesis protein moeb.
GN MOEB OR STM0845.
OS Salmomella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmomella.
OX NCBI_Taxid=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2;
RA Wong K.K., Kwan H.S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SCSG1412 / ATCC 700720;
RA MEDLINE=21534948; PubMed=11677609;
RA McCelland M., Sanderson K.E., Splieth J., Clifton S.W., Latreille P.,
RA Courtney L., Portwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmomella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
CC (MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING
CC FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U53178; AAA96530.1; -
DR EMBL: AE008735; AAL19781.1; -
DR StGene: SG10665; moeb.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; Thif_domain.
DR Pfam: PF00899; Thif. 1.
KW Molybdenum cofactor biosynthesis; Complete proteome.
FT CONFLICT 9 9 M->I (IN REF. 1).
FT CONFLICT 38 38 G->R (IN REF. 1).
FT CONFLICT 117 117 S->A (IN REF. 1).
FT CONFLICT 169 169 N->T (IN REF. 1).
FT CONFLICT 221 221 G->E (IN REF. 1).
SQ SEQUENCE 249 AA; 26903 MW; 0F0050831D557AD2 CRC64;

Query Match 2.58; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNRQI 185
DB 10 LRYNRQI 16
|||||

```

```

RESULT 38
MOEB_SALTY STANDARD; PRT; 249 AA.
AC 056067;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)

```

```

RESULT 39
KDTX_SERMA
ID KDTX_SERMA STANDARD; PRT; 257 AA.
AC 054435;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Lipopolysaccharide core biosynthesis glycosyl transferase kdtx
DE (EC 2.4.1.14).
GN KDTX.
OS Serratia marcescens.

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N28B;
RX MEDLINE=96422003; PubMed=8824620;
RA Guasch J.F., Pique N., Climent N., Ferrer S., Merlino S., Rubires X.,
RA Tomas J.M., Regue M.;
RT "Cloning and characterization of two Serratia marcescens genes
RT involved in core lipopolysaccharide biosynthesis.";
RL J. Bacteriol. 178:5741-5747(1996)
CC -1- PATHWAY: Lipopolysaccharide core biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2. WAE/KDTX
CC SUPERFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U52844; AAC4443.1; -
DR InterPro: IPR001173; Glycosyltransf_2.
DR Pfam: PF00535; Glycosyltransf_2; 1.
SQ LipoPolysaccharide biosynthesis; Transferase; Glycosyltransferase.
KW SEQUENCE 257 AA; 29233 MW; D40D7B57E002F990 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 257;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GLLAVV 77
|111111|
DB 229 GLLAVV 235

RESULT 40
TNFC_MARMO
ID TNFC_MARMO STANDARD; PRT; 310 AA.
AC 09JUN10: 09JUN11;
DT 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
DE necrosis factor ligand superfamily member 3).
GN LTB OR TNFSF3 OR TNFC.
OS Marmota monax (Woodchuck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
OC Marmota.
OX NCBI_TaxID=9995;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20184748; PubMed=10721723;
RA Li D.H., Haveli E.A., Brown C.L., Cullen J.M.;
RT "Woodchuck lymphotoxin-alpha, -beta and tumor necrosis factor genes:
RT structure, characterization and biological activity.";
RL Gene 242:295-305(2000).
CC -1- FUNCTION: Cytokine that binds to LTB/TNFSF3. May play a specific
CC role in immune response regulation. Provides the membrane anchor
CC for the attachment of the heterotrimeric complex to the cell
CC surface.
CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or
CC (less prevalent) two LTA and one LTB subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF096268; AAF3486.1; -
DR EMBL: AF095587; AAF3486.1; -
DR HSSP: P01374; 1TNF.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF; 1.
DR PRINTS: PR01234; TNFCROSISFCT.
DR PRODOM: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor.
FT DOMAIN 1 27 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 28 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 49 310 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 280 280 D -> H (IN REF. 1; AAF34865).
SQ SEQUENCE 310 AA; 32644 MW; 73B354EFC8B333BE CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 310;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GLYLVC 200
|111111|
DB 181 GLYLVC 187

```

Search completed: March 31, 2003, 14:14:17  
Job time : 30 secs

125  
Briedel 5/15/03  
Page 1

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 13:51:31 ; Search time 76 Seconds  
(without alignments)  
769,966 Million cell updates/sec

Title: US-09-245-198a-4  
Perfect score: 284  
Sequence: 1 MSLLDFEISARLRPLRSLG.....PNAHLKARPLTYFGLEQYH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 65 summaries

Database :

- 1: SPTEMBL\_21.\*
- 2: sp.archaea.\*
- 3: sp.bacteria.\*
- 4: sp.fungi.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.podent.\*
- 12: sp.virus.\*
- 13: sp.vertebrate.\*
- 14: sp.unclassified.\*
- 15: sp.rvivirus.\*
- 16: sp.bacteriap.\*
- 17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	3.5	111	16	08X4J8
2	9	3.2	142	16	0981J0
3	8	2.8	51	4	09BX52
4	8	2.8	143	17	09HST7
5	8	2.8	151	10	09SD11
6	8	2.8	184	5	09VV70
7	8	2.8	197	16	09WYU0
8	8	2.8	199	4	09BU11
9	8	2.8	211	5	08SXH4
10	8	2.8	278	16	09ZG99
11	8	2.8	279	11	09D378
12	8	2.8	279	11	09CPR8
13	8	2.8	306	5	08SZB8
14	8	2.8	339	17	058554
15	8	2.8	342	4	099908
16	8	2.8	343	4	09BPV2

# ALIGNMENTS

RESULT 1	ID	08X4J8	PRELIMINARY;	PRT;	111 AA.
AC	08X4J8	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Hypothetical protein z3516.				
GN	z3516.				
OS	Escherichia coli O157:H7.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC	Escherichia.				
OX	NCBI_TaxID=83334;				
RN	11				
RP	SEQUENCE FROM N.A.				
RC	SPRAIN=O157:H7 / EDL933 / ATCC 700927;				
RX	MEDLINE=21074935; PubMed=11206551;				
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,				

17	8	2.8	370	17	08TU08	08TU08 methanopyru
18	8	2.8	372	16	053860	053860 mycobacteri
19	8	2.8	377	16	09KRT3	09KRT3 streptomyc
20	8	2.8	387	10	09SHD8	09SHD8 arabidopsis
21	8	2.8	397	16	080820	080820 agrobacteri
22	8	2.8	435	16	08XS70	08XS70 ralsionia s
23	8	2.8	443	4	09UFM5	09UFM5 homo sapien
24	8	2.8	454	5	09VNP0	09VNP0 drosophila
25	8	2.8	455	16	08ZCV8	08ZCV8 yeisina pe
26	8	2.8	471	16	09RR18	09RR18 delnoccoc
27	8	2.8	472	4	096N66	096N66 homo sapien
28	8	2.8	473	11	09CV76	09CV76 mus musculu
29	8	2.8	473	11	08R1P9	08R1P9 mus musculu
30	8	2.8	522	10	09L1W0	09L1W0 oryza sativ
31	8	2.8	564	12	09Y1T09	09Y1T09 influenzavi
32	8	2.8	568	12	08QPL0	08QPL0 influenza a
33	8	2.8	656	5	09N8H2	09N8H2 trypanosoma
34	8	2.8	656	5	09N8U8	09N8U8 trypanosoma
35	7	2.5	1523	5	09N8U8	09N8U8 trypanosoma
36	7	2.5	24	5	P83215	P83215 octopus vul
37	7	2.5	30	5	P83217	P83217 octopus vul
38	7	2.5	32	6	09BER6	09BER6 capra hircu
39	7	2.5	35	2	09ZG35	09ZG35 chlamydia t
40	7	2.5	54	4	016193	016193 homo sapien
41	7	2.5	67	2	08RQ54	08RQ54 seriatia ma
42	7	2.5	69	5	09VMM9	09VMM9 drosophila
43	7	2.5	86	16	09AAG4	09AAG4 caulobacter
44	7	2.5	94	4	09P074	09P074 homo sapien
45	7	2.5	99	2	P97199	P97199 escherichia
46	7	2.5	100	5	08WQ08	08WQ08 eupyryma sc
47	7	2.5	100	12	08QRH6	08QRH6 heparitis c
48	7	2.5	101	15	08UTL1	08UTL1 human immun
49	7	2.5	108	9	038214	038214 bacterioph
50	7	2.5	115	16	09HY60	09HY60 pseudomonas
51	7	2.5	118	2	0939P7	0939P7 pseudomonas
52	7	2.5	118	16	0928G5	0928G5 listeria in
53	7	2.5	119	10	094LX6	094LX6 euphorbia t
54	7	2.5	121	10	08W0M1	08W0M1 oryza sativ
55	7	2.5	123	16	08UC82	08UC82 agrobacteri
56	7	2.5	125	5	09U1P6	09U1P6 caenorhabd1
57	7	2.5	126	1	0977X7	0977X7 methanococ
58	7	2.5	133	15	091121	091121 human immun
59	7	2.5	135	4	09BT89	09BT89 homo sapien
60	7	2.5	145	4	09UMX4	09UMX4 homo sapien
61	7	2.5	150	4	09Y413	09Y413 homo sapien
62	7	2.5	156	10	09ACR8	09ACR8 oryza sativ
63	7	2.5	157	10	049178	049178 oryza sativ
64	7	2.5	157	16	09RYB7	09RYB7 delnoccoc
65	7	2.5	159	10	08W0T0	08W0T0 sorghum b1c

RA Posfai G., Hackett J., Klink S., Boufin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,  
RA Apodaca J., Aantaahtaman T.S., Lin J., Ten G., Schwartz D.C.,  
RA Welch R.A., Blatner F.R.,  
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
RL Nature 409:529-533(2001).  
DR EMBL: AE005458: AAG57389.1; -.  
DR InterPro: IPR000620; DUF6.  
DR Pfam: PF00892; DUF6; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 111 AA; 12165 MW; 7CEFC93D786CD759 CRC64;

Query Match 3.5%; Score 10; DB 16; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.13;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGALACGL 72  
DB 40 LGALACGL 49

RESULT 2  
Q981J0 PRELIMINARY; PRT; 142 AA.  
ID Q981J0:  
AC Q981J0: 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE Probable transport protein.  
GN MLR2380.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFE303099;  
RX MEDLINE=21082930; Pubmed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP002999; BAB49526.1; -.  
KW Complete Proteome.  
SQ SEQUENCE 142 AA; 14884 MW; 0DCA7842CB5A5E6F CRC64;

Query Match 3.2%; Score 9; DB 16; Length 142;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 TALVPLAL 61  
DB 122 TALVPLAL 130

RESULT 3  
Q9BX52 PRELIMINARY; PRT; 51 AA.  
ID Q9BX52:  
AC Q9BX52: 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE DJ81IC21.1 (Novel protein similar to gamma-glutamyl  
DE transpeptidase-related protein (GGT-Rel)) (Fragment).  
GN DJ81IC21.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;

RN [1]  
RP SEQUENCE FROM N.A.  
RA Corby N.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL133466; CAC34607.1; -.  
FT NON\_TER 51 51  
SQ SEQUENCE 51 AA; 5124 MW; 31980CCE0451EF8 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 51;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
DB 7 LGIGLALA 14

RESULT 4  
Q9HST7 PRELIMINARY; PRT; 143 AA.  
ID Q9HST7:  
AC Q9HST7: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Vn90080H.  
GN Vn90080H.  
OS Halobacterium sp. (Strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; Pubmed=11016950.  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
RA Shukla H.D., Laskey S.R., Balliga N.S., Thorsson V., Shroga J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,  
RA Lelhauser B., Keller K., Cruz R., Danon M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isehaberger T.A., Beck R.F., Pohlshroder M., Spidlich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL: AE004976; AAG18715.1; -.  
KW Complete Proteome.  
SQ SEQUENCE 143 AA; 15648 MW; 45466E328EF3468 CRC64;

Query Match 2.8%; Score 8; DB 17; Length 143;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
DB 55 LGIGLALA 62

RESULT 5  
Q9SD11 PRELIMINARY; PRT; 151 AA.  
ID Q9SD11:  
AC Q9SD11: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Hypothetical protein (OSJNB40036E02.6 protein) (B1085F09.2  
DE protein).  
GN B1085F09.2.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriatoidae; Oryzae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.

RC SFRAIN-CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC  
 clone: P000310.";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SFRAIN-CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC  
 clone: OSCNB0036E02.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SFRAIN-CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC  
 clone: B1085F09.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP000815; BAA87834.1; -;  
 DR EMBL: AP002862; BAB1732.1; -;  
 DR EMBL: AP003103; BAB44106.1; -;  
 SQ SEQUENCE 151 AA; 16632 MW; EC68451ECA2BD71D CRC64;

Query Match 2.8%; Score 8; DB 10; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRGGE 50  
 Db 131 RRRGRGGE 138

RESULT 6  
 O9VV70 PRELIMINARY; PRT; 184 AA.  
 AC O9VV70;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE CG13033 protein.  
 GN CG13033.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amandlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cwley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Folsler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003527; AAF9452.1; -;  
 DR FLYBase: FBgn0036638; CG13033.  
 DR InterPro: IPR004019; YLP\_motif.  
 DR Pfam: PF02757; YLP\_5.  
 SQ SEQUENCE 184 AA; 21022 MW; 31976AE350DDC447 CRC64;

Query Match 2.8%; Score 8; DB 5; Length 184;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLAAVSL 79  
 Db 16 LLAAVSL 23

RESULT 7  
 O9WYU0 PRELIMINARY; PRT; 197 AA.  
 AC O9WYU0;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein TM0469.  
 GN TM0469.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
 OX NCBI\_TaxID=2336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SFRAIN-MSB / DSM 3109;  
 RX MEDLINE=99287316; PubMed=10360571;  
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
 genome sequence of Thermotoga maritima.";  
 RL Nature 399:323-329(1999).  
 DR EMBL: AE001724; AAD35553.1; -;  
 DR TIGR: TM0469;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 197 AA; 22919 MW; 41E2C8E3C09180BC CRC64;

Query Match 2.8%; Score 8; DB 16; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SLDPFEIS 9  
 Db 135 SLDPFEIS 142

RESULT 8  
 O9BU11 PRELIMINARY; PRT; 199 AA.  
 ID O9BU11

```

AC 09BU11;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Hypothetical 22.8 kda protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN 11;
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002598; AAH02598.1; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
KM Hypothetical protein.
SQ SEQUENCE 199 AA; 22750 MW; E5DAC747E7E8F06B CRC64;

Query Match
Best Local Similarity 2.8%; Score 8; DB 4; Length 199;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLAL 67
DB 167 ALGLAL 174

RESULT 9
OBSXH4 PRELIMINARY; PRT; 211 AA.
ID 08SXH4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE RES0345D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokslein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rudin G.M.,
RA Ceiliker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY089634; AAL90372.1; -.
SQ SEQUENCE 211 AA; 23780 MW; 82FF4983E91F510A CRC64;

Query Match
Best Local Similarity 2.8%; Score 8; DB 5; Length 211;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 LLLAVSL 79
DB 9 LLLAVSL 16

RESULT 10
O9ZG99 PRELIMINARY; PRT; 278 AA.
ID 09ZG99;
AC 09ZG99;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Transmembrane protein AMPE.
GN AMPE OR PA4521.

```

```

OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RA Langae Y.T., Dargis M., Huletsky A.;
RT "An ampD gene in Pseudomonas aeruginosa encodes a negative regulator
of ampC beta-lactamase expression.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE-20437337; PubMed-10984043;
RA Stoyer C.K., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Coltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Bardo L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AF082575; AAC96784.1; -.
DR EMBL; AE004866; AAC07909.1; -.
KM Transmembrane; Complete proteome.
SQ SEQUENCE 278 AA; 30793 MW; C623F1AB0691CFF CRC64;

Query Match
Best Local Similarity 2.8%; Score 8; DB 16; Length 278;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LGLLAV 77
DB 47 LGLLAV 54

RESULT 11
O9D378 PRELIMINARY; PRT; 279 AA.
ID 09D378;
AC 09D378;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE 5730494616Rik protein.
GN 5730494616Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-MEDULLA OBLONGATA;
RX MEDLINE-21085660; PubMed-11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Arakawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuwaji P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittlaker C., Wilmink L.,
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;

```

RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK018250; BAB31133.1; -  
 DR MGD: MGI:1913897; 5730494G16Rik.  
 DR InterPro: IPR002190; MAGE.  
 DR Pfam: PF01454; MAGE.1.  
 SQ SEQUENCE 279 AA; 31474 MW; 5E243590A99F15F0 CRC64;

Query Match 2.8%; Score 8; DB 11; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 80 GSRASLSA 87  
 Db 41 GSRASLSA 48

RESULT 12  
 Q9CPR8 PRELIMINARY; PRT; 279 AA.

AC Q9CPR8;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DE 5730494G16Rik protein (MAGE-91).  
 GN 5730494G16Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN=CS7BL/6J; TISSUE=EMBRYO, AND EMBRYONIC STEM CELLS;  
 RX MEDLINE=21085560; Pubmed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kodota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Ginstreich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA Auguer P.H., Chomez P.M., De Backer O.R., Bertrand M.J.M.;  
 RT "Ten new murine members of the MAGE gene family."  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK017727; BAB30899.1; -  
 DR EMBL: AK010294; BAB26830.1; -  
 DR EMBL: AF319979; AAK01207.1; -  
 DR MGD: MGI:1913897; 5730494G16Rik.  
 DR InterPro: IPR002190; MAGE.  
 DR Pfam: PF01454; MAGE.1.  
 SQ SEQUENCE 279 AA; 31460 MW; FE2435919B63160 CRC64;

Query Match 2.8%; Score 8; DB 11; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 80 GSRASLSA 87  
 Db 41 GSRASLSA 48

Db 41 GSRASLSA 48

RESULT 13

ID Q852B8 PRELIMINARY; PRT; 306 AA.  
 AC Q852B8;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE RE07882P.  
 GN CG1169.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyroidae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunoo J., Paclebb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celinker S.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY070982; AAL48604.1; -  
 SQ SEQUENCE 306 AA; 34083 MW; 32B69371475448F9 CRC64;

Query Match 2.8%; Score 8; DB 5; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 56 LVPALGL 63  
 Db 170 LVPALGL 177

RESULT 14  
 ID 058554 PRELIMINARY; PRT; 339 AA.  
 AC 058554;  
 DT 01-AUG-1998 (TREMblrel. 07, Created)  
 DT 01-JAN-1999 (TREMblrel. 09, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Hypothetical protein PH0824.  
 GN PH0824.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=53953;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3;  
 RX MEDLINE=98344137; Pubmed=9679194;  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohkura Y.,  
 RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL: AP000003; BAA29917.1; -  
 DR InterPro: IPR002657; Bileac/Na\_smprtr.  
 DR Pfam: PF01758; SBF.1.  
 KW Hypothetical protein, Complete proteome

Query Match 2.8%; Score 8; DB 17; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 31;

Oy 80 GSRASLSA 87  
 Db 41 GSRASLSA 48

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 70 LGLLAVV 77  
|||||||  
DB 111 LGLLAVV 118

RESULT 15  
O99908 PRELIMINARY: PRT: 342 AA.  
AC O99908:  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE BBL protein.  
GN BBL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-96273128; PubMed-8702217;  
RX Fukunaga-Johnson N., Lee S.W., Liebert M., Grossman H.B.;  
RT "Molecular analysis of a gene, BBL, overexpressed in bladder and  
RT breast carcinoma."  
RL Anticancer Res. 16:1085-1090(1996).  
DR EMBL: S82470; AAB37433.1; -;  
DR InterPro: IPR004299; MBOAT\_fam.  
DR Pfam: PF03062; MBOAT; 1  
SQ SEQUENCE 342 AA; 38163 MW; 2B479E8CFB91C CRC64;

Query Match 2.8%; Score 8; DB 4; Length 342;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 ALGGLAL 67  
|||||||  
DB 310 ALGGLAL 317

RESULT 16  
O9BPV2 PRELIMINARY: PRT: 343 AA.  
AC O9BPV2:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Unknown (protein for MGC:4221) (protein for MGC:2029).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-KIDNEY, AND COLON;  
RA Strusberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC003164; AAH03164.1; -;  
DR EMBL: BC002513; AAH02512.1; -;  
DR InterPro: IPR004299; MBOAT\_fam.  
DR Pfam: PF03062; MBOAT; 1  
SQ SEQUENCE 343 AA; 38727 MW; F71E7DBF7ABD9BB7 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 343;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 ALGGLAL 67  
|||||||  
DB 311 ALGGLAL 318

RESULT 17  
O8TU08 PRELIMINARY: PRT: 370 AA.  
AC O8TU08:  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Permease subunit of a ABC-type transport system involved in  
DE lipoprotein release.  
GN MK1655.  
OS Methanopyrus kandleri.  
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
OC Methanopyrus.  
OX NCBI\_TaxID=2320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AV19 / DSM 6324 / JCM 9639;  
RX MEDLINE-21927647; PubMed-11930014;  
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,  
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
RA Natile D.A., Rogozin I.B., Talusov R.L., Wolf Y.I., Stetter K.O.,  
RA Mal'kh A.G., Koonin E.V., Kozlovskiy S.A.;  
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
RT and monophyly of archaeal methanogens."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
DR EMBL: AE010455; AAM02868.1; -;  
KW Complete proteome.  
SQ SEQUENCE 370 AA; 39411 MW; B0762EAE1E5A644E CRC64;

Query Match 2.8%; Score 8; DB 17; Length 370;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 61 LGGLALA 68  
|||||||  
DB 336 LGGLALA 343

RESULT 18  
O53860 PRELIMINARY: PRT: 372 AA.  
AC O53860:  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Hypothetical protein cyam3.  
GN CYM3 OR RV0848 OR MTV043.41.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV;  
RX MEDLINE-98295987; PubMed-9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltham D., Gentles S., Hamlin N., Holroyd S.,  
RA Horsby T., Jagels K., Krogh A., McLaren J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).  
DR EMBL: AL022004; CAA17654.1; -;  
DR HSSP: P35520; LJBQ.  
DR Tuberculist: RV0848; -;  
DR InterPro: IPR001926; B6\_enzyme\_beta.  
DR Pfam: PF00291; PALP; 1  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 372 AA; 40118 MW; 927386BE1DE5FB6C CRC64;



Query Match 2.8%; Score 8; DB 16; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGLGLALA 68  
 |||||  
 Db 96 LGLGLALA 103

RESULT 19  
 O9KVT3

PRELIMINARY; PRT; 377 AA.

AC O9KVT3;  
 DT 01-OCT-2000 (TREMUREL. 15, Created)  
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)  
 DE Putative integral membrane protein.  
 GN SC05682 OR SC5H4.06C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;

RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)."  
 RL Nature 417:141-147(2002).  
 DR EMBL; AL355913; CAB91118.1; -  
 SQ SEQUENCE 377 AA; 37614 MW; A35DA0437F044C6 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LAUGLGLA 66  
 |||||  
 Db 327 LAUGLGLA 334

RESULT 20  
 O9SHD8

PRELIMINARY; PRT; 387 AA.

AC O9SHD8;  
 DT 01-MAY-2000 (TREMUREL. 13, Created)  
 DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)  
 DE At2g45002 protein (At2g45000/TIAP1.20) (Hypothetical 40.6 kDa  
 protein).  
 GN At2g45000 OR At2g45005.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;

RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RX MEDLINE=20083487; Pubmed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Beil C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,

RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Coppenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Selzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana."  
 RL Nature 402:761-768(1999).

RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RN  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Lim J., Meyers M.C., Banh J.,  
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RT "Arabidopsis cDNA clones."  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]

RN  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Arabidopsis full length cDNA clones."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC007659; AAD32835.1; -  
 DR EMBL; AY074646; AAL69462.1; -  
 DR EMBL; AY080620; AAL69303.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 387 AA; 40584 MW; AF6C6B3BAC9BF69A CRC64;

Query Match 2.8%; Score 8; DB 10; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 ELVAEED 100  
 |||||  
 Db 370 ELVAEED 377

RESULT 21  
 O8U820

PRELIMINARY; PRT; 397 AA.

AC O8U820;  
 DT 01-JUN-2002 (TREMUREL. 21, Created)  
 DT 01-JUN-2002 (TREMUREL. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)  
 DE Hypothetical protein AtU3948.

GN AtU3948 OR AGR\_L1608.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=176299;

RN  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21608550; Pubmed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kitayarin T., Levy R., Li M.-J., McClelland E., Palmeri A.,  
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,

RA Chunmley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
RA Neater E.W.;  
RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*  
RT C58.";  
RL Science 294:2317-2323(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21608551; PubMed-11743194;  
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
RA Wollam C., Allinger M., Doughy D., Scott C., Iappas C., Markelz B.,  
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,  
RA Cielo C., Slater S.;  
RT "Genome sequence of the plant pathogen and biotechnology agent  
RT *Agrobacterium tumefaciens* C58.";  
RL Science 294:2323-2328(2001).  
DR EMBL: AE009325; AAL44750.1; ALT\_INIT.  
DR EMBL: AE008289; AAK89478.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 397 AA; 41708 MW; 700748E32A46AE86 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 397;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 251 LALRPGSS 258  
|||||  
Db 341 LALRPGSS 348

## RESULT 22

O8XS70 PRELIMINARY; PRT; 435 AA.  
AC O8XS70;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Probable transmembrane protein.  
GN RSP0611 OR RS03756.  
OS *Ralstonia solanacearum* (Pseudomonas solanacearum).  
OC Plasmid megaplasmid.  
OC Bacteria; Proteobacteria; beta subdivision; *Ralstonia* group;  
OC *Ralstonia*.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-GM11000;  
RX MEDLINE-21681879; PubMed-11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
RA Chaudler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Sigler P., Trebault P., Whalen M., Wincker P., Levy M.,  
RA Weisenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*,"  
RL Nature 415:497-502(2002).  
DR EMBL: AL646079; CAD17762.1; -.  
DR InterPro: IPR001564; NDK.  
DR PROSITE: PS00469; NDK\_KINASES; UNKNOWN\_1.  
KW Plasmid; Complete proteome.  
SQ SEQUENCE 435 AA; 47048 MW; CCD859D9C54DDB5A CRC64;

Query Match 2.8%; Score 8; DB 16; Length 435;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 217 LVDGVIAL 224  
|||||  
Db 358 LVDGVIAL 365

## RESULT 23

O9UFM5 PRELIMINARY; PRT; 443 AA.  
AC O9UFM5;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical 47.4 kDa protein.  
GN DKFZP5660011.  
OS *Homo sapiens* (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RL Blum H., Baerends S., Mewes H.W., Gassenhuber J., Wiemann S.;  
RT Submitted (SEP-1999) to the EMBL/Genbank/DBD databases.  
DR EMBL: AL117414; CAB55910.1; -.  
DR MEROPS: T03.002; -.  
DR InterPro: IPR00101; G4lutnsptdase.  
DR Pfam: PF01019; G\_glu\_transpept. 1.  
DR PRINTS: PR01210; GGTANSPTASE.  
KW Hypothetical protein.  
SQ SEQUENCE 443 AA; 47446 MW; 3D75E9DB0B265971 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 443;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 61 LGIGIALA 68  
|||||  
Db 14 LGIGIALA 21

## RESULT 24

O9VNP0 PRELIMINARY; PRT; 454 AA.  
AC O9VNP0;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE CG1169 protein.  
GN CG1169.  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Ball J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spredley A.C., Stapleton M., Strong R., Sun E.,  
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glibe R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003600; AAF5189.1; -;  
 DR FlyBase: FBgn0037428; CG1169.  
 DR FLYBASE: 454 AA; 51320 MW; A75AAD97E716573 CRC64;  
 SQ SEQUENCE

Query Match 2.8%; Score 8; DB 5; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 56 LVPLALGL 63  
 |||||  
 Db 170 LVPLALGL 177

RESULT 25  
 082CV8 PRELIMINARY; PRT; 465 AA.  
 AC 082CV8;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Putative membrane protein.  
 GN YEGB OR YPO2850.  
 OS *Yersinia pestis*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC *Yersinia*.  
 OX NCBI\_TaxID=632;  
 OX [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-CO-92 / BIOVAR ORIENTALIS;  
 RX MDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titchall R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Feltrielli T., Hamlin N., Holtroyd S., Jagers K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skellern J., Stevens K., Whitehead S., Barrell B.G.;  
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague.";  
 RL Nature 413:523-527(2001).  
 DR EMBL: A414154; CAC92102.1; -;  
 DR InterPro: IPR003662; sub\_transporter.  
 DR Pfam: PF00083; sugar\_tr; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 465 AA; 50176 MW; 0CC273F10B3F5BD CRC64;

Query Match 2.8%; Score 8; DB 16; Length 465;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 72 LLAVALSL 79  
 |||||  
 Db 335 LLAVALSL 342

RESULT 26  
 O9RR18

ID O9RR18 PRELIMINARY; PRT; 471 AA.  
 AC O9RR18;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE Transpot protein, putative.  
 GN DR2502.  
 OS *Deinococcus radiodurans*.  
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RI;  
 RX MDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA McFitt K.S., Qin H., Jiang L., Pamphili W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Usterback T., Zaleski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium *Deinococcus*  
 RT *radiodurans* RI.";  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AE002079; AAF12043.1; -;  
 DR TIGR: DR2502; -;  
 KW Complete proteome.  
 SQ SEQUENCE 471 AA; 47974 MW; 96B2BEBF6E445D27 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 471;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 ALGLGLAL 67  
 |||||  
 Db 366 ALGLGLAL 373

RESULT 27  
 096N66 PRELIMINARY; PRT; 472 AA.  
 ID 096N66;  
 AC 096N66;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE CNA FLJ31346 fts, clone MESAN1000180, highly similar to BBI-malignant  
 DE cell expression-enhanced gene/tumor progression-enhanced gene.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Kawal-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,  
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,  
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,  
 RA Isegaki T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK055908; BAB71043.1; -;  
 DR InterPro: IPR004299; MBOAT\_fam.  
 DR Pfam: PF03062; MBOAT; 1.  
 SQ SEQUENCE 472 AA; 52774 MW; EA721998043F9EBD CRC64;

Query Match 2.8%; Score 8; DB 4; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALACLGAL 67  
 |||||||  
 DB 440 ALACLGAL 447

## RESULT 28

OYCY76 PRELIMINARY: PRT: 473 AA.  
 AC O9CY76:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE 5730589L02RIK protein;  
 GN 5730589L02RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,  
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 Alkawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 Fleischmann W., Gaasterland T., Glasl C., King B., Kochia H.,  
 Kuhl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,  
 Blake J., Botfield D., Boujunga N., Carninci P., de Bonaldo M.F.,  
 Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,  
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,  
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuk S.,  
 Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection";  
 RT Nature 409:685-690(2001).  
 RL EMBL: AK019981; BAB31950.1; -;  
 DR MGD: MGI:1924832; 5730589L02RIK.  
 DR InterPro: IPR004299; MBOAT\_fam.  
 DR Pfam: PF03062; MBOAT\_1.  
 SQ SEQUENCE 473 AA; 53504 MW; CEGF8E93C3D01C4F CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 72  
 |||||||  
 DB 436 LALACLG 443

## RESULT 29

OBR1P9 PRELIMINARY: PRT: 473 AA.  
 AC OBR1P9:  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE RIKEN CDNA 5730589L02 gene.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 DR Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC023417; AAH23417.1; -;  
 SQ SEQUENCE 473 AA; 53382 MW; DAA1FE0DA78013EA CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 72  
 |||||||  
 DB 436 LALACLG 443

## RESULT 30

OYLIW0 PRELIMINARY: PRT: 522 AA.  
 AC OYLIW0:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE Similar to an Arabidopsis thaliana chromosome BAC genomic  
 DE sequence.  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Euphorbiaceae; Oryzaeae; Oryza.  
 OX NCB1\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hsing Y.C., Chow T., Chen C., Wu H., Chu M., Chao Y., Liu S.;  
 RT "Oryza sativa PAC P0699E04 genomics sequence, complete sequence."  
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP001111; BAA90509.1; -;  
 SQ SEQUENCE 522 AA; 54697 MW; 21C6BAD2441B56BF CRC64;

Query Match 2.8%; Score 8; DB 10; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 46;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRGE 50  
 |||||||  
 DB 415 RRRGRGE 422

## RESULT 31

O9YTU9 PRELIMINARY: PRT: 564 AA.  
 AC O9YTU9:  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hemagglutinin H5 (Fragment).  
 OS Influenzavirus A.  
 CC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 CC Influenza A viruses.  
 OX NCB1\_TaxID=11320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A/DUCK/PORTSDAM/2216-4/84;  
 RX MEDLINE=99099002; PubMed=9882316;  
 RA Matrosovich M., Zhou N., Kawaoka Y., Webster R.;  
 RT "The surface glycoproteins of H5 influenza viruses isolated from  
 humans, chickens, and wild aquatic birds have distinguishable  
 properties.";  
 RT J. Virol. 73:1146-1155(1999).  
 CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO  
 CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS  
 CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.  
 DR EMBL: AF082041; AAD13573.1; -;  
 DR HSSP: P03437; 1HTM.  
 DR InterPro: IPR001364; Hemagglutn.  
 DR Pfam: PF00509; Hemagglutinin; 1.

DR PRINTS: PR00329; HEMAGGLUTININ2.  
DR Prodom: PD000225; Hemagglutin. 1.  
KW Envelope protein; Glycoprotein; Hemagglutinin.  
FT NON\_TER 564 564  
SQ SEQUENCE 564 AA; 63562 MW; B317179A7F3E6F98 CRC64;

Query Match 2.8%; Score 8; DB 12; Length 564;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVSL 79  
DB 6 LLLAVVSL 13

## RESULT 32

OS Influenza A virus (A/Goose/Hong Kong/3014.8/2000(H5N1)).  
OC Influenza A viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OX NCBI\_TaxID=176675;  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=A/GOOSE/HONG KONG/3014.8/2000(H5N1);  
RX MEDLINE=21874832; PubMed=11878904;  
RA Guan Y., Peiris M., Kong K.F., Dyrting K.C., Ellis T.M., Sit T.,  
Zhang L.J., Shortridge K.F.;  
RT "H5N1 Influenza Viruses Isolated from Geese in Southeastern China:  
Evidence for Genetic Reassortment and Interspecies Transmission to  
Ducks.";  
RT Virology 292:16-23(2002).  
DR EMBL: AY059482; AL31388.1; -.  
FT NON\_TER 568 568  
SQ SEQUENCE 568 AA; 64281 MW; DB0A4CFE034F1769 CRC64;

Query Match 2.8%; Score 8; DB 12; Length 568;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVSL 79  
DB 6 LLLAVVSL 13

## RESULT 33

ID Q9N8H2 PRELIMINARY; PRT; 656 AA.  
AC Q9N8H2;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
DE Possible hypothetical 61.9 kDa protein.  
GN CHRI.338.  
OS Trypanosoma brucei.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5691;  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=TREU927;  
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
Gerrard C., Rajandream M.A., Barrell B.G.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL359782; CAB95571.1; -.  
FT NON\_TER 656 656  
SQ SEQUENCE 656 AA; 72138 MW; CBAC892D25937FAD CRC64;

Query Match 2.8%; Score 8; DB 5; Length 656;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 BELVAEED 100  
DB 454 BELVAEED 461

## RESULT 34

ID Q9N8U8 PRELIMINARY; PRT; 1523 AA.  
AC Q9N8U8;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Possibly hypothetical protein 85.6 kD.  
GN CHRI.139.  
OS Trypanosoma brucei.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxID=5691;  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=TREU927;  
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
Gerrard C., Rajandream M.A., Barrell B.G.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL359782; CAB95435.1; -.  
FT NON\_TER 1523 1523  
SQ SEQUENCE 1523 AA; 168322 MW; 077BDC751CD1E5A CRC64;

Query Match 2.8%; Score 8; DB 5; Length 1523;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 VDGVLALR 225  
DB 382 VDGVLALR 389

## RESULT 35

ID P83215 PRELIMINARY; PRT; 24 AA.  
AC P83215;  
DT 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Sperm protamine P3 (P03) (Fragment).  
OS Octopus vulgaris (Octopus).  
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
OC Incirrata; Octopodidae; Octopus.  
OX NCBI\_TaxID=6645;  
RN  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
RC TISSUE=SPERM;  
RA Gimenez-Bonate P., Ribes E., Buena C., Sautiere P., Kouach M.,  
Aurio J., Kasinsky H.E., Chiva M.;  
RT "Chromatin remodeling and protamines during spermiogenesis of Octopus  
vulgaris (Cephalopoda).";  
RL J. Exp. Zool. 0:0-0(2001).  
DE -1- FUNCTION: PROTEIN SUBSTRATE FOR HISTONES IN THE CHROMATIN OF  
SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- TISSUE SPECIFICITY: TESTIS.  
CC -1- MASS SPECTROMETRY: MM=4389; METHOD=ELECTROSPRAY.  
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
DNA condensation; Nuclear protein.  
FT DOMAIN 1 16  
FT NON\_TER 24 24  
SQ SEQUENCE 24 AA; 3381 MW; 308E90ED9D2C9C9C CRC64;

Query Match 2.5%; Score 7; DB 5; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 1111111  
 Db 11 RRRGRG 17

RESULT 36  
 P83217 PRELIMINARY; PRT; 30 AA.  
 AC P83217:  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Sperm protamine p5 (Pos).  
 OS Octopus vulgaris (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incurtata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=6645;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC TISSUE=SPERM;  
 RA Gimenez-Bonafe P., Ribes E., Buesa C., Sautiere P., Kouach M.,  
 RA Ausio J., Kasinsky H.E., Chiva M.;  
 RT "Chromatin remodeling and protamines during spermiogenesis of Octopus  
 RT vulgaris (Cephalopoda).";  
 RL J. Exp. Zool. 0:0-0(2001).  
 CC -1- FUNCTION: PROTAGENES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- MASS SPECTROMETRY: MW=3941; METHOD=ELECTROSPRAY.  
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW DNA condensation; Nuclear protein.  
 FT DOMAIN 2 15 POLY-ARG.  
 FT DOMAIN 17 26 POLY-ARG.  
 SQ SEQUENCE 30 AA; 3943 MW; 14F1BC7E4D277049 CRC64;

Query Match 2.5%; Score 7; DB 5; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 1111111  
 Db 5 RRRGRG 11

RESULT 37  
 Q9BEF6 PRELIMINARY; PRT; 32 AA.  
 AC Q9BEF6:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Beta-lactoglobulin precursor (Fragment).  
 GN BETA-LACTOGLOBULIN.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA D'Andrea M., Pilla F., Graziano M.;  
 RT "A new polymorphism of goat beta-lactoglobulin proximal promoter  
 RT region.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ292058; CAC27455.1; -.  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.

FT NON-TER 32 32  
 SQ SEQUENCE 32 AA; 3372 MW; 0C56BD579B30C190 CRC64;  
 Query Match 2.5%; Score 7; DB 6; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGLALAC 69  
 1111111  
 Db 8 LGLALAC 14

RESULT 38  
 Q9ZG35 PRELIMINARY; PRT; 35 AA.  
 AC Q9ZG35:  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE Hypothetical 3.5 kDa protein (Fragment).  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=L2 434B;  
 RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
 RT "Gene identification of Chlamydia trachomatis by random DNA  
 RT sequencing.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF087333; AAD04107.1; -.  
 KW Hypothetical protein.  
 FT NON-TER 1 1  
 FT NON-TER 35 35  
 SQ SEQUENCE 35 AA; 3519 MW; 4668B72BAF28F7D CRC64;

Query Match 2.5%; Score 7; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 NSSSPLR 180  
 1111111  
 Db 8 NSSSPLR 14

RESULT 39  
 O16193 PRELIMINARY; PRT; 54 AA.  
 AC O16193:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Orf2 5' to PD-ECGF/TP protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94312438; PubMed=8038210;  
 RA Usuki K., Gomez L.J., Wernstedt C., Moren A., Miyazono K.,  
 RA Claesson-Welsh L., Heldin C.H.;  
 RT "Structural properties of 3.0 kb and 3.2 kb transcripts encoding  
 RT platelet-derived endothelial cell growth factor/thymidine  
 RT phosphorylase in A431 cells.";  
 RL Biochim. Biophys. Acta 1222:411-414(1994).  
 DR EMBL; S72487; AAD14107.1; -.  
 SQ SEQUENCE 54 AA; 5808 MW; 827925FBA70CB222 CRC64;

Query Match 2.5%; Score 7; DB 4; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 RPRRSAP 131  
|||||||  
Db 17 RPRRSAP 23

## RESULT 40

Q8R0S4 PRELIMINARY; PRT; 67 AA.  
ID Q8R0S4  
AC Q8R0S4;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Hypothetical 7.8 kDa protein.  
GN YHCR.  
OS Serratia marcescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Serratia.  
OX NCBI\_TaxID=615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Baba M., Midorikawa Y., Nakagawa Y., Fujita M., Matsuyama T.;  
RT "Serratia marcescens and Escherichia coli genes controlling  
temperature-dependent production of structurally unrelated secondary  
metabolites such as prodigiosin and serrawettin.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB080601; BAB85653.1; -.  
KM Hypothetical protein  
SQ SEQUENCE 67 AA: 7791 MM; 47B01A87E69AC2A2 CRC64;  
  
Query Match 2.5%; Score 7; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 55 LVPPLAL 61  
|||||||  
Db 21 LVPPLAL 27

Search completed: March 31, 2003, 14:15:52  
Job time : 88 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:45 ; Search time 53.5639 Seconds

(without alignments)  
706.506 Million cell updates/sec

Title: US-09-245-198a-4

Sequence: 1 MSLLDPEISARRLPLPRSLG.....PNAHLKAPFLRYFGLFQVH 284

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

A.Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1444	100.0	284	19	AAW47525
2	1268	87.8	249	20	AAI09369
3	1268	87.8	249	21	AAW07526
4	1268	87.8	249	21	AAW53338
5	1268	87.8	249	23	AAU086129
6	1265	87.6	249	19	AAW29745
7	1265	87.6	249	22	AAE00891
8	1066	73.8	273	22	AAU03499
9	1062	73.5	208	20	AAW93590
10	1020	70.6	225	19	AAW47524

11	1020	70.6	225	21	AAW07527
12	968	67.0	211	20	AAW93591
13	792	54.8	189	19	AAW29746
14	792	54.8	189	22	AAE00892
15	761	52.7	146	22	AAE00895
16	116	8.0	325	22	ABW67553
17	116	8.0	409	23	AAU77718
18	108.5	7.5	211	21	AAW58216
19	108.5	7.5	260	21	AAW58215
20	106.5	7.4	254	16	AAW64190
21	106.5	7.4	254	18	AAW26657
22	106.5	7.4	254	23	ABW59533
23	106	7.3	1428	21	AAW97033
24	105.5	7.3	406	23	AAU77717
25	104.5	7.2	779	23	ABW07845
26	104	7.2	409	23	AAU77716
27	104	7.2	1323	15	AAW5248
28	102	7.1	256	22	AAW25657
29	100.5	6.9	647	17	AAW04327
30	100	6.9	220	22	AAW62340
31	99	6.9	574	21	AAW97032
32	99	6.9	1008	22	AAW78891
33	99	6.9	1020	22	AAW79875
34	96.5	6.7	1097	22	ABG25655
35	96.5	6.7	1631	22	ABG22481
36	96	6.6	234	22	AAW62339
37	95	6.6	240	23	AAE13680
38	95	6.6	294	19	AAW6956
39	95	6.6	294	19	AAW68292
40	95	6.6	294	22	AAE08737
41	95	6.6	294	22	AAE04425
42	95	6.6	294	22	AAE01992
43	94.5	6.5	1006	22	ABG21178
44	94	6.5	876	22	ABG00217
45	93.5	6.5	409	22	AAW00076

## ALIGNMENTS

RESULT 1	
AAW47525	
ID	AAW47525 standard; Protein: 284 AA.
XX	
AC	AAW47525;
XX	
DT	21-JUL-1998 (first entry)
XX	
DE	Homo sapiens tumour necrosis factor related ligand (TRELL).
XX	
KW	TRELL, tumour necrosis factor related ligand; tnfr; treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection.
XX	
OS	Homo sapiens.
XX	
PN	WO9805783-A1.
PD	12-FEB-1998.
XX	
PF	07-AUG-1997; 97WO-US13945.
XX	
PR	18-MAR-1997; 97US-0040820.
PR	07-AUG-1996; 96US-0023541.
XX	
PR	18-OCT-1996; 96US-0028515.
PA	(BIOJ ) BIOGEN INC.
XX	(UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX	
PI	Browning JL, Chicheportliche Y;
XX	
DR	WPI, 1998-145619/13.
DR	N-PSDB; AAV18600.

```
XX Tumour necrosis factor related ligand - useful for, e.g. treating
PT cancer, auto-immune disease and immune responses to tissue grafts
XX
XX Claim 12; Pages 50-51; 69pp; English.
PS
XX
XX The sequence is that of human tumour necrosis factor related
CC ligand (TRELL). TRELL or active fragments can be included with a
CC carrier in pharmaceutical compositions to treat cancer, autoimmune
CC diseases or immune responses to tissue grafts, or to stimulate or
CC suppress the immune system. It is useful to screen for TRELL
CC receptors, by labelling with a detectable label and screening
CC compositions for binding. Agents interfering with TRELL-receptor
CC binding can also be screened for, can then be administered,
CC optionally with interferon- gamma, to induce cell death or
CC treat, suppress or alter immune responses (especially involving human
CC adenocarcinoma cells) involving a signal pathway between TRELL and its
CC receptor. It's coding sequence can be used in gene therapy for
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
CC autoimmune and inflammatory diseases or inherited genetic disorders,
CC by introducing into cells, and expressing, therapeutically effective
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
CC It may also be of use in the preparation of prepare probes for
CC screening natural/synthetic DNAs for TRELL-encoding sequences
CC and for antisense therapy.
CC
XX
SQ Sequence 284 AA:
Query Match 100.0%; Score 1444; DB 19; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.2e-129;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSLLDFEISARLPPLPRSGISGRDGAIVRQAPPAARRRSGRRGEPGTALLVPLA 60
DB 1 MSLLDFEISARLPPLPRSGISGRDGAIVRQAPPAARRRSGRRGEPGTALLVPLA 60
OY 61 LGLGLAACLGLLAVVSLGSRASLSAOPAEELVAEEDOPSELNPTQESSQDAPPL 120
DB 61 LGLGLAACLGLLAVVSLGSRASLSAOPAEELVAEEDOPSELNPTQESSQDAPPL 120
OY 121 NMLVPRRSAPKGRKTRARRAIAHYEVHPRPGODGAQGVGTVSGWEARINSSPLR 180
DB 121 NMLVPRRSAPKGRKTRARRAIAHYEVHPRPGODGAQGVGTVSGWEARINSSPLR 180
OY 181 YNRQIGEFIVTRAGLYLYTCQVHFDEGKAVYLKLDLVGVALRCLLEFSATTAASLGP 240
DB 181 YNRQIGEFIVTRAGLYLYTCQVHFDEGKAVYLKLDLVGVALRCLLEFSATTAASLGP 240
OY 241 QURLCOVSGLLALRPSSSLRIRTLPMANHLKAAPFLTYEGLFOVH 284
DB 241 QURLCOVSGLLALRPSSSLRIRTLPMANHLKAAPFLTYEGLFOVH 284
RESULT 2
AAY09369
ID AAY09369 standard; Protein: 249 AA.
AC
XX AAY09369;
XX
XX 15-JUL-1999 (first entry)
XX
XX Human tumour necrosis factor Apo-3 ligand protein sequence.
XX
XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
XX NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
XX cancer.
XX
XX Homo sapiens.
XX
XX OS
XX PN MO9919490-A1.
XX PD 22-APR-1999.
XX
```

```
PF 09-OCT-1998; 98WO-US21407.
XX
XX 17-DEC-1997; 97US-0069862.
XX 10-OCT-1997; 97US-0062037.
XX
XX (GENTH) GENENTECH INC.
XX
XX Ashkenazi AJ, Marsters SA, Pittl R;
XX WPI; 1999-287982/24.
XX N-PSDB; AAX56000.
XX
XX New human Apo-3 ligand (a tumour necrosis factor) homologue
XX
XX Claim 1; Fig 1; 74pp; English.
XX
XX The present sequence represents a human tumour necrosis factor (TNF)
XX and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
XX cytostatic activity. Apo-3 ligand can be used to induce apoptosis in
XX mammalian cancer cells, to induce NF-kappaB-dependent transcription and
XX to induce JNK/SAPK-dependent responses in mammalian cells.
XX
XX
SQ Sequence 249 AA:
Query Match 87.8%; Score 1268; DB 20; Length 249;
Best Local Similarity 100.0%; Pred. No. 6.6e-113;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 36 MAARSORRRGRGEPGTALLVPLALGLALACGLLAVVSLGSRASLSAOPAEEL 95
DB 1 MAARSORRRGRGEPGTALLVPLALGLALACGLLAVVSLGSRASLSAOPAEEL 95
OY 96 VAEEDDPSELNPTQESSQDAPPLNMLVPRRSAPKGRKTRARRAIAHYEVHPRGOD 155
DB 96 VAEEDDPSELNPTQESSQDAPPLNMLVPRRSAPKGRKTRARRAIAHYEVHPRGOD 120
OY 61 VAEEEDDPSELNPTQESSQDAPPLNMLVPRRSAPKGRKTRARRAIAHYEVHPRGOD 120
DB 61 VAEEEDDPSELNPTQESSQDAPPLNMLVPRRSAPKGRKTRARRAIAHYEVHPRGOD 120
OY 156 GAQAGVDGTVSGWEARINSSPLRKNRQIGEFIVTRAGLYLYTCQVHFDEGKAVYLKLD 215
DB 156 GAQAGVDGTVSGWEARINSSPLRKNRQIGEFIVTRAGLYLYTCQVHFDEGKAVYLKLD 180
OY 121 GAQAGVDGTVSGWEARINSSPLRKNRQIGEFIVTRAGLYLYTCQVHFDEGKAVYLKLD 180
DB 121 GAQAGVDGTVSGWEARINSSPLRKNRQIGEFIVTRAGLYLYTCQVHFDEGKAVYLKLD 180
OY 216 LLYDGVVALRCLLEFSATTAASSLGPQLRICOVSGLLALRPSSSLRIRTLPMANHLKAAPFL 275
DB 181 LLYDGVVALRCLLEFSATTAASSLGPQLRICOVSGLLALRPSSSLRIRTLPMANHLKAAPFL 240
OY 276 TYEGLFOVH 284
DB 241 TYEGLFOVH 249
RESULT 3
AAB07526
ID AAB07526 standard; Protein: 249 AA.
AC
XX AAB07526;
XX
XX 20-OCT-2000 (first entry)
XX
XX Amino acid sequence of a soluble recombinant human TWEAK protein.
XX
XX TWEAK protein; immunological disorder; immune response; inflammation;
XX TWEAK blocking agent; autoimmune disease; organ transplant rejection;
XX Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.
XX
XX Homo sapiens.
XX
XX OS
XX PN WO200042073-A1.
XX PD 20-JUL-2000.
XX
XX 14-JAN-2000; 2000WO-US01044.
XX
XX 15-JAN-1999; 99US-0116168.
XX
```



Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 36 MAARSRQRRGRGEPGTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPQDEL 95
DB 1 MAARSRQRRGRGEPGTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPQDEL 60
OY 96 VAEEDODPSLNPQTEESODPAFLNRLVPRRSAPKGGKTTARRAIAAHYEVHPPGCD 155
DB 61 VAEEDODPSLNPQTEESODPAFLNRLVPRRSAPKGGKTTARRAIAAHYEVHPPGCD 120
OY 156 GAQAGVDGTVSGMEARINSSPLRYNRQIGFEITYRAGLYLYLCOVHDEGKAVYTKLD 215
DB 121 GAQAGVDGTVSGMEARINSSPLRYNRQIGFEITYRAGLYLYLCOVHDEGKAVYTKLD 180
OY 216 LLYDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHLKAAPFL 275
DB 181 LLYDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHLKAAPFL 240
OY 276 TYFGLFQVH 284
DB 241 TYFGLFQVH 249
```

## RESULT 5

AAU86129  
ID AAU86129 standard: Protein; 249 AA.

```
AC AAU86129;
DT 15-JUL-2002 (first entry)
DE Human PRO207 polypeptide.
XX
KW Human; PRO: benign tumour; malignant tumour; lymphoid malignancy;
KW leukaemia; neuronal disorder; stromal disorder; blastocoealic disorder;
KW inflammatory disorder; immune disorder; angiogenic disorder;
KW cytoskeletal; neuroprotective.
```

XX Homo sapiens.

OS WO200153486-A1.

PD 26-JUL-2001.

PF 11-FEB-2000; 2000WO-US03565.

```
PR 08-MAR-1999; 99WO-US05028.
PR 11-MAR-1999; 99US-123972P.
PR 11-MAY-1999; 99US-133459P.
PR 02-JUN-1999; 99WO-US12252.
PR 22-JUN-1999; 99US-140650P.
PR 22-JUN-1999; 99US-140653P.
PR 20-JUL-1999; 99US-144758P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 17-AUG-1999; 99US-149395P.
PR 31-AUG-1999; 99US-151689P.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 05-JAN-2000; 2000WO-US00219.
```

XX (GETH ) GENENTECH INC.

PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AJ, Hillan KJ;

PI Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;

PI Watanabe CK, Wood WI;

XX WPI: 2002-205567/26.

DR N-PSDB: ABK40255.

```
XX Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukaemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
XX
PS Claim 61; Fig 4; 302pp; English.
```

XX The present invention relates to the isolation of novel human PRO  
CC polypeptides and the polynucleotide sequences encoding them. The  
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are  
CC useful for treating benign or malignant tumours (e.g. renal, kidney,  
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other  
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,  
CC macrophagal, stromal and blastocoealic disorders, inflammatory, immune  
CC and angiogenic disorders. The polynucleotide sequences are also  
CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO  
CC polypeptides of the invention.

XX Sequence 249 AA:

Query Match 87.8%; Score 1268; DB 23; Length 249;  
Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 36 MAARSRQRRGRGEPGTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPQDEL 95
DB 1 MAARSRQRRGRGEPGTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPQDEL 60
OY 96 VAEEDODPSLNPQTEESODPAFLNRLVPRRSAPKGGKTTARRAIAAHYEVHPPGCD 155
DB 61 VAEEDODPSLNPQTEESODPAFLNRLVPRRSAPKGGKTTARRAIAAHYEVHPPGCD 120
OY 156 GAQAGVDGTVSGMEARINSSPLRYNRQIGFEITYRAGLYLYLCOVHDEGKAVYTKLD 215
DB 121 GAQAGVDGTVSGMEARINSSPLRYNRQIGFEITYRAGLYLYLCOVHDEGKAVYTKLD 180
OY 216 LLYDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHLKAAPFL 275
DB 181 LLYDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMHLKAAPFL 240
OY 276 TYFGLFQVH 284
DB 241 TYFGLFQVH 249
```

## RESULT 6

AAW29745  
ID AAW29745 standard: Protein; 249 AA.

AC AAW29745;

DT 27-OCT-1998 (first entry)

DE TNF related endothelium proliferative agent protein.

KW TNF; endothelium proliferative agent; TRBPA; wound healing; cancer;

KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.

XX Homo sapiens.

OS WO9835061-A2.

PD 13-AUG-1998.

PF 12-FEB-1998; 98WO-US02859.

PR 10-FEB-1998; 98US-0021706.

PR 12-FEB-1997; 97US-0798692.

PA (ABBO ) ABBOTT LAB.

PI Wiley SR;

XX

DR WPI: 1998-447255/38.  
 DR N-PSDB; AAV47613.  
 XX Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 XX  
 XX  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 XX  
 SQ Sequence 249 AA:  
 Query Match 87.6%; Score 1265; DB 19; Length 249;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;  
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 36 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVNSIGSRASISAOEPAGEEL 95  
 DB 1 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVNSIGSRASISAOEPAGEEL 60  
 QY 96 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 155  
 DB 61 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 120  
 QY 156 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYLCQVHFDGKAVYTKLD 215  
 DB 121 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYLCQVHFDGKAVYTKLD 180  
 QY 216 LTVDGVLALRCLEEFSAATASISLGPQLRLCOVSGLLALRPSSLRIRTLPAHILKAAPFL 275  
 DB 181 LTVDGVLALRCLEEFSAATASISLGPQLRLCOVSGLLALRPSSLRIRTLPAHILKAAPFL 240  
 QY 276 TYFGLFOVH 284  
 DB 241 TYFGLFOVH 249  
 RESULT 7  
 AAE00891  
 ID AAE00891 standard; Protein; 249 AA.  
 XX  
 AC AAE00891;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 XX  
 XX Human; tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KM TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KM grafting; vulnerary.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 98..249  
 FT /label= Extracellular\_domain  
 XX  
 PN US6207642-B1.  
 XX

PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PT Wiley SR;  
 XX  
 DR WPI: 2001-280760/29.  
 DR N-PSDB; AAD04350.  
 XX  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 XX  
 PS Claim 1; Column 75-76; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is clone ID #690050 human TREPA.  
 CC  
 XX  
 SQ Sequence 249 AA:  
 Query Match 87.6%; Score 1265; DB 22; Length 249;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;  
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 36 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVNSIGSRASISAOEPAGEEL 95  
 DB 1 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVNSIGSRASISAOEPAGEEL 60  
 QY 96 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 155  
 DB 61 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 120  
 QY 156 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYLCQVHFDGKAVYTKLD 215  
 DB 121 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLYLCQVHFDGKAVYTKLD 180  
 QY 216 LTVDGVLALRCLEEFSAATASISLGPQLRLCOVSGLLALRPSSLRIRTLPAHILKAAPFL 275  
 DB 181 LTVDGVLALRCLEEFSAATASISLGPQLRLCOVSGLLALRPSSLRIRTLPAHILKAAPFL 240  
 QY 276 TYFGLFOVH 284  
 DB 241 TYFGLFOVH 249  
 RESULT 8  
 AAU03499  
 ID AAU03499 standard; Protein; 273 AA.  
 XX  
 AC AAU03499;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE TWEAK extracellular domain-containing fusion protein.  
 XX  
 XX TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KM rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
 KM corneal graft neovascularisation; psoriasis; metastatic condition;  
 KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KM

KM preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KM scleroderma; vascular adhesion; telangiectasia; ischemia; human;  
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KM peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAK;  
 KM fusion protein.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200145730-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 XX 19-DEC-2000; 2000WO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMMV ) IMMUNEX CORP.  
 XX  
 PI Willey SR;  
 PI  
 DR WPI: 2001-417975/44.  
 DR N-PSDB; AAS03964.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor -  
 XX  
 PS Example 1; Page 41; 46pp; English.  
 XX  
 CC The sequence represents a fusion protein encoded by the  
 CC expression vector pDC409-L2-TWEAK. The fusion protein comprises a  
 CC growth hormone leader, a leucine zipper multimerisation domain, and  
 CC the extracellular domain of human TWEAK. The fusion protein was  
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 CC clones from a COS cell human cDNA library. The TWEAK protein is  
 CC a member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and  
 CC develop TWEAKR agonists and antagonists for the modulation of  
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
 CC The disorders mediated by angiogenesis include ocular disorders  
 CC characterised by ocular neovascularisation such as diabetic retinopathy,  
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 CC retrolental fibroplasia, rubecosis, uveitis, macular degeneration and  
 CC corneal graft neovascularisation, and inflammatory diseases such as  
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
 CC malignant and metastatic conditions such as sarcomas and carcinomas,  
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischemia.  
 CC  
 XX  
 XX Sequence 273 AA:  
 SO  
 Query Match 73.8%; Score 1066; DB 22; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-93;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 247 SLRITLPMWHLKAPFLTYFGLFQVH 273  
 ||||||||||||||||||||||||||||  
 RESULT 9  
 AAW93590  
 ID AAW93590 standard; Protein; 208 AA.  
 XX  
 AC AAW93590;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Human TNFR3 protein.  
 XX  
 KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KM developmental abnormality; gestational abnormality; prostate cancer;  
 KM APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KM apoptosis; human.  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 PM WO9911791-A2.  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 XX (UNIV ) UNIV WASHINGTON.  
 PA  
 PI Chaudhary PM;  
 PI  
 DR WPI: 1999-205191/17.  
 DR N-PSDB; AAX23424.  
 XX  
 PR New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PR useful for diagnosis and treatment of prostate cancer and  
 PR developmental or gestational abnormalities  
 XX  
 PS Claim 40; Fig 13A; 156pp; English.  
 PS  
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 CC  
 XX  
 SO Sequence 208 AA:  
 Query Match 73.5%; Score 1062; DB 20; Length 208;  
 Best Local Similarity 99.5%; Pred. No. 2.6e-93;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 VSLSGRASLSAEPQAEELVAEEDDPSELNPQTEESQDPAPFLNRLVRRPRAKGRKT 60  
 QY 137 RARRATAHAHEVHPRPQDDAAGVDSVSGWEARINSSSPLRYNROIGEFYTRAGLY 196  
 DB 61 RARRATAHAHEVHPRPQDDAAGVDSVSGWEARINSSSPLRYNROIGEFYTRAGLY 120  
 QY 197 YLYCOVHFDEGKAVYLYKLDLVGVLALRCLFEFSATAASSLGQPLRLCOVSGGLALRPG 256  
 DB 121 YLYCOVHFDEGKAVYLYKLDLVGVLALRCLFEFSATAASSLGQPLRLCOVSGGLALRPG 180  
 QY 257 SSLRIRLTPMAHLKAAFLTYFGFLQVH 284  
 DB 181 SSLRIRLTPMAHLKAAFLTYFGFLQVH 208

RESULT 10  
 AAM47524  
 ID AAM47524 standard; Protein; 225 AA.  
 XX  
 AC AAM47524;  
 XX  
 DT 21-JUL-1998 (first entry)  
 XX  
 DE Mus musculus tumour necrosis factor related ligand (TRELL).  
 XX  
 KW TRELL; tumour necrosis factor related ligand; tnf; treatment;  
 KM cancer; autoimmune disease; immune system; stimulation; suppression;  
 KM graft rejection.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..21  
 FT /note= "hydrophobic, transmembrane domain"  
 XX  
 PN WO9805783-A1.  
 PD 12-FEB-1998.  
 XX  
 PF 07-AUG-1997; 97MO-US13945.  
 XX  
 PR 18-MAR-1997; 97US-0040820.  
 PR 07-AUG-1996; 96US-0023541.  
 PR 18-OCT-1996; 96US-0028515.  
 XX  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.  
 XX  
 PI Browning JL, Chicheportiche Y;  
 XX  
 DR WPI; 1998-145619/13.  
 DR N-PSDB; AAV18599.  
 XX  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 12; Pages 48-50; 69pp; English.

CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of prepare probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 CC  
 XX Sequence 225 AA;  
 SQ  
 Query Match 70.6%; Score 1020; DB 19; Length 225;  
 Best Local Similarity 88.8%; Pred. No. 2, 9e-89;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 61 LGIGLALACGLLLAVVSLGRASLSAEPQAEELVAEEDDPSELNPQTEESQDPAPFL 120  
 DB 2 LSLGLALACGLLLVAVVSLGSMATLSAEPQAEELVAEEDDPSELNPQTEESQDPAPFL 61  
 QY 121 NRIYRPRRSAPKGRKTARRAIAHAYEHVHPRPQDDAAGVDSVSGWEARINSSSPLR 180  
 DB 62 EQLVRRPRRSAPKGRKTARRAIAHAYEHVHPRPQDDAAGVDSVSGWEARINSSSPLR 121  
 QY 181 YNROIGEFYTRAGLYLYLYCOVHFDEGKAVYLYKLDLVGVLALRCLFEFSATAASSLGP 240  
 DB 122 YDRQIGEFYTRAGLYLYLYCOVHFDEGKAVYLYKLDLVGVLALRCLFEFSATAASSLGP 181  
 QY 241 QLRLCOVSGGLALRPGSSLRIRLTPMAHLKAAFLTYFGFLQVH 284  
 DB 182 QLRLCOVSGGLALRPGSSLRIRLTPMAHLKAAFLTYFGFLQVH 225

RESULT 11  
 AAB07527  
 ID AAB07527 standard; protein; 225 AA.  
 XX  
 AC AAB07527;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soluble recombinant murine TWEAK protein.  
 XX  
 KW TWEAK protein; immunological disorder; immune response; inflammation;  
 KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
 KM Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200042073-A1.  
 PD 20-JUL-2000.  
 XX  
 PF 14-JAN-2000; 2000MO-US01044.  
 XX  
 PR 15-JAN-1999; 99US-0116168.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PI Renneert P;  
 XX  
 DR WPI; 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease -  
 XX  
 PS Disclosure; Fig 1; 45pp; English.

CC The present sequence represents a TWEAK protein. The specification  
 CC describes a method for preventing or treating an immunological  
 CC disorder and/or inhibiting an immune response in an animal. The  
 CC method comprises administering a TWEAK blocking agent. The method may  
 CC be used for preventing and treating immune disorders associated with  
 CC inappropriate expression and/or activity of TWEAK. These disorders  
 CC include autoimmune diseases, acute and chronic inflammation, organ  
 CC transplant rejection, Graft-versus-host disease (GVHD), lymphoid cell  
 CC malignancies, septic and other forms of shock, loss of immune





CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
CC promote wound healing or tissue grafting, by promoting vascularisation,  
CC also to induce apoptosis for treating cancer and eliminating autoreactive  
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
CC TREPA peptides can also be used to target cytotoxic agents or for  
CC affinity isolation of the corresponding receptor, the nucleic acid for  
CC which can be used to transform tumour cells to render them more  
CC responsive to TREPA and to screen for TREPA mimics.  
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
CC vascularisation), inflammation or a wide range of autoimmune conditions,  
CC conditions involving abnormal stimulation of epithelial cells (e.g.  
CC atherosclerosis), for birth control (inhibiting ovulation and placental  
CC formation) or other angiogenic conditions (e.g. ulcers).  
CC  
XX  
SQ Sequence 189 AA;  
Query Match 54.8%; Score 792; DB 19; Length 189;  
Best Local Similarity 99.3%; Pred. No. 1.4e-67;  
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 132 KGRTRARRAIAAHYEVHPRGQDAGVGTGSGWEARINSSPLRYNRQIGEFIVT 191  
DB 37 KGRTRARRGIAAHYEVHPRGQDAGVGTGSGWEARINSSPLRYNRQIGEFIVT 96  
QY 192 RAGLYTYLCQVHFDEGKAVYTKLDLVDGVLAALRCLEEFSAATASSLGAPQLRQVSGLL 251  
DB 97 RAGLYTYLCQVHFDEGKAVYTKLDLVDGVLAALRCLEEFSAATASSLGAPQLRQVSGLL 156  
QY 252 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFQVH 284  
DB 157 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFQVH 189  
RESULT 14  
AAE00892  
ID AAE00892 standard; Protein: 189 AA.  
XX  
AC AAE00892;  
XX  
DT 04-JUL-2001 (first entry)  
XX  
DE Human U14flag TREPA soluble construct.  
XX  
KM Human; tumour necrosis factor: TNF; angiogenesis; wound healing; tumour;  
KM TREPA; TNF related endothelium proliferative agent; metastasis; grafting;  
KM vulnery; HUVEC; human umbilical vein endothelial cell; U14flag.  
XX  
OS Homo sapiens.  
XX  
PN US6207642-B1.  
XX  
PD 27-MAR-2001.  
XX  
PF 26-JUN-1998; 98US-0105343.  
XX  
PR 12-FEB-1997; 97US-0798692.  
PR 10-FEB-1998; 98US-0021706.  
XX  
PA (ABBO ) ABBOTT LAB.  
XX  
PI WILEY SR;  
XX  
DR WPI; 2001-280760/29.  
XX  
PT Inducing angiogenesis in mammal at desired sites for promoting wound  
PT healing, by administering soluble fragment of extracellular domain of  
PT tumor necrosis factor related endothelium proliferative agent protein  
PT  
XX  
PS Example 2; Column 75-78; 53pp; English.  
XX  
CC The present invention relates to extracellular signal molecules,

CC particularly members of tumour necrosis factor (TNF) family molecules  
CC designated as TREPA (TNF related endothelium proliferative agent).  
CC Soluble biologically active TREPA are used to treat TREPA-associated  
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
CC in human for promoting wound healing and for vascularising grafted  
CC tissue for successful grafting and to promote tissue grafts. The present  
CC amino acid sequence is human U14flag TREPA soluble construct. This  
CC sequence which is a biologically active molecule is capable of inducing  
CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.  
XX  
SQ Sequence 189 AA;  
Query Match 54.8%; Score 792; DB 22; Length 189;  
Best Local Similarity 99.3%; Pred. No. 1.4e-67;  
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 132 KGRTRARRAIAAHYEVHPRGQDAGVGTGSGWEARINSSPLRYNRQIGEFIVT 191  
DB 37 KGRTRARRGIAAHYEVHPRGQDAGVGTGSGWEARINSSPLRYNRQIGEFIVT 96  
QY 192 RAGLYTYLCQVHFDEGKAVYTKLDLVDGVLAALRCLEEFSAATASSLGAPQLRQVSGLL 251  
DB 97 RAGLYTYLCQVHFDEGKAVYTKLDLVDGVLAALRCLEEFSAATASSLGAPQLRQVSGLL 156  
QY 252 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFQVH 284  
DB 157 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFQVH 189  
RESULT 15  
AAE00895  
ID AAE00895 standard; Protein: 146 AA.  
XX  
AC AAE00895;  
XX  
DT 04-JUL-2001 (first entry)  
XX  
DE Human TREPA (TNF related endothelium proliferative agent) fragment.  
XX  
KM Human; tumour necrosis factor: TNF; angiogenesis; wound healing;  
KM TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
KM grafting; vulnery.  
XX  
OS Homo sapiens.  
XX  
PN US6207642-B1.  
XX  
PD 27-MAR-2001.  
XX  
PF 26-JUN-1998; 98US-0105343.  
XX  
PR 12-FEB-1997; 97US-0798692.  
PR 10-FEB-1998; 98US-0021706.  
XX  
PA (ABBO ) ABBOTT LAB.  
XX  
PI WILEY SR;  
XX  
DR WPI; 2001-280760/29.  
XX  
PT Inducing angiogenesis in mammal at desired sites for promoting wound  
PT healing, by administering soluble fragment of extracellular domain of  
PT tumor necrosis factor related endothelium proliferative agent protein  
PT  
XX  
PS Example 14; Fig 1; 53pp; English.  
XX  
CC The present invention relates to extracellular signal molecules,  
CC particularly members of tumour necrosis factor (TNF) family molecules  
CC designated as TREPA (TNF related endothelium proliferative agent).  
CC Soluble biologically active TREPA are used to treat TREPA-associated  
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
CC in human for promoting wound healing and for vascularising grafted

CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is human TREPA fragment.

xx  
 SQ Sequence 146 Aa;

Query Match 52.7%; Score 761; DB 22; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-65;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 RRAIAAHVEVHPRPGODGAQAGVDGTVSGWEARINSSPLRYNROIGETVTRAGIYYL 198  
 |||  
 Db 1 RRAIAAHVEVHPRPGODGAQAGVDGTVSGWEARINSSPLRYNROIGETVTRAGIYYL 60  
 QY 199 YCQVHFDEGKAVYLLKLDLLVDGVLALRCLLEFSATAASSLGQRLCQVSGLLALRPGSS 258  
 |||  
 Db 61 YCQVHFDEGKAVYLLKLDLLVDGVLALRCLLEFSATAASSLGQRLCQVSGLLALRPGSS 120  
 QY 259 LRIRTLPMWHLKAAAPFLTYFGLFQVH 284  
 |||  
 Db 121 LRIRTLPMWHLKAAAPFLTYFGLFQVH 146

Search completed: March 31, 2003, 07:44:12  
 Job time : 55.5639 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 00:53:03 ; Search time 3581.52 Seconds  
(without alignments)  
2307.731 Million cell updates/sec

Title: US-09-245-198a-4  
Perfect score: 1444  
Sequence: 1 MSLLDFEISARRLPLPRSLG.....FWAHLKAAFLTYGLFQVH 284

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPRO/spool/US09245198/tunat\_24032003.163553.25950/app.query.fasta.1.846  
-DB=genbml -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human4.0.cdi -LIST=45  
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09245198.qcgn.1.1.3077.ernat.24032003.163553.25950 -NCPU=6 -ICPU=3  
-NO.XLPHY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genbml:\*  
1: gb\_ba:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pla:\*  
35: em\_htg\_rpd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1366	94.6	1353	6	AX201324
2	1366	94.6	1368	9	AF055872
3	1331.5	92.2	1651	9	BC019047
4	1298	89.9	1306	9	AF030099
5	1265	87.6	1236	6	AR140407
6	1066	73.8	898	6	AX180714
7	1020	70.6	1168	10	AF030100
8	589.5	40.8	177703	2	AC016876
9	589.5	40.8	215795	2	AC127470
10	557	38.6	177555	2	AC130192
11	555.5	38.5	161428	2	AC126925
12	535.5	38.5	203083	2	AC069459
13	555.5	38.5	234182	10	AL603707
14	495	34.3	153553	2	AC126921
15	489.5	33.9	138792	2	AC119115
16	385.5	26.7	184026	2	AC098923
17	214	14.8	203281	2	AC126237
18	145.5	10.1	187835	8	AP002482
19	140	9.7	42210	1	SC1C
20	138.5	9.6	190050	1	AL646059
21	136.5	9.5	242930	2	AC105470
22	135	9.3	6397	1	AF002222
23	135	9.3	10728	1	AE004651
24	134	9.3	10965	1	AE004740
25	130	9.0	40352	1	SC61A
26	129	8.9	5751	1	SGDNASTRS
27	129	8.9	16185	1	SCC46
28	129	8.9	25459	1	SGAJ6985
29	129	8.9	32941	1	SCM10
30	129	8.9	43961	1	SCM1
31	128	8.9	137833	9	CNS01DV6
32	127	8.8	194304	9	AC016708
33	126.5	8.8	207050	9	AL646063
34	125.5	8.7	999	9	HSBARP11
35	125.5	8.7	28795	1	AVINIRC
36	125.5	8.7	87144	8	AP003822
37	125.5	8.7	153595	2	AP004272
38	125	8.7	30590	1	SC4G2
39	125	8.7	191221	2	AP005292
40	124.5	8.6	36734	1	SC6G10
41	124.5	8.6	98504	2	AP003876
42	124.5	8.6	107132	2	AP003911
43	124.5	8.6	154746	14	HSV2HG52
44	124.5	8.6	176580	2	AP003634
45	124	8.6	3957	6	A45258

RESULT 1

## ALIGNMENTS

[illegible]

QY	217	LeuValAspArglyValLeuAlaAlaLeuArgGlySerLeuGluLysuphSerAlaThrAlaAlaSer	236
Db	601	TTTTGGTGGATGGTGTCTGGCCCTGGCGTCCCTGGAGGAATTCAGCCACATGGCGGAGGT	660
QY	237	SerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGly	256
Db	661	TCCCTCGGGGCCCCACACTCCGGCTCTGGCCAGGTGTCTGGCGCTGTGGCCCTCGGGCGAAGG	720
QY	257	SerSerLeuArgIleArgThrLeuProTrpAlaHisLeuAlaHisAlaAlaProPheLeuThr	276
Db	721	TCCCTCCCTGGCGGATCGGACCCCTCCCTGGGGCCCATCTCAAGGCTGCCCCCTCTCTCAC	780
QY	277	TyrPheGlyLeuPheGlnValHis	284
Db	781	TACTTCGGACCTCTTCACAGTTTCAC	804
RESULT 2			
AF055872			
LOCUS	AF055872	1368 bp	mRNA linear PRI 04-MAY-1998
DEFINITION	Homo sapiens Apoc3/DR3 ligand (Apoc3L) mRNA, complete cds.		
ACCESSION	AF055872		
VERSION	AF055872.1	GI:3108230	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1368) Masters,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
TITLE	Identification of a ligand for the death-domain-containing receptor Apoc3		
JOURNAL	Curr. Biol. 8 (9), 525-528 (1998)		
MEDLINE	98228355		
PUBMED	9560343		
REFERENCE	2 (bases 1 to 1368) Masters,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
AUTHORS	Direct Submission Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way, South San Francisco, CA 94080, USA		
TITLE	Location/Qualifiers		
JOURNAL	1..1368		
FEATURES	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17p13" 1..1368 /gene="APO3L" 58..807 /gene="APO3L" /function="binds to the death-domain-containing receptor Apoc3/DR3" /note="TNF homolog" /codon_start=1 /product="Apoc3/DR3 ligand" /protein_id="AAC39724.1" /db_xref="GI:3108231" /translation="MAAKRSQRRRRRGEGFTALLPVALGCLALACGLLAVSLGSRSLSAQEPQAELEVEDDPSSELNPTGESDPAFLNLTVPRRRSPGGRYTRARRAIAAHVEYHPRQGDGAQAGVDGVSQMEBARINSSSPLRYNROIGEIFYTRAGLYIKLCVHPDEGKAVLYLKLDLVGVALRCLEEFSSATASSISGPDRLRCQVSGLLALRPSSLRIRLRLPMNHAKAPFLIYFGLFQYH"		
BASE COUNT	272 a 443 c 389 g 264 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	4..71e-63	Length:	1368
Score:	1366.00	Matches:	268
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	94.60%	Indels:	0

DB:	9	Gaps:	0
US-09-245-198a-4 (1-284) x AF055872 (1-1368)			
QY	17	ArgSerLeuGlySerArgAspGlyAlaValArgGlnAlaGlnProProAlaProMet	36
Db	1	CGATCCCTCGGGGTCGCGGATGGGGGGGGGTGAGGAGCAGACCCCGCCCGCCATG	60
QY	37	AlaAlaArgArgSerGlnArgArgArgGlyArgArgGlyGluProGlyTThAlaLeu	56
Db	61	GCCGCCGCTGGAGCCAGAGCGGAGGGGGCGCGGGGGAGCGCGCCCTGTG	120
QY	57	ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaLeuAlaLeuAlaVal	76
Db	121	GTCCCGCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG	180
QY	77	ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal	96
Db	181	GTCAGTTGGGAGCCCGGAGCATGCTGTGCGCCAGAGAGCTGCCAGAGAGCTGTG	240
QY	97	AlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro	116
Db	241	GCAAGAGAGGACCAAGGACCCCTCGAATCGAATCCAGACAGAAAGCCAGAGATCCT	300
QY	117	AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgLysThr	136
Db	301	GCGCCTTTCCTGAACCCGACTAGTTCGGCTCGCAGAAAGTGCACCTMAAGCCGGAAACA	360
QY	137	ArgAlaArgArgAlaAlaAlaAlaHisTyrGluValHisProArgProGlyGlnAspGly	156
Db	361	CGGGCTCGAAGAGGATCGCAGCCATTATGAGTTCATCCAGACCTCGAGAGACGGA	420
QY	157	AlaGlnAlaGlyValaLeuAspGlyThrValSerGlyTrrGluGlnAlaArgLysSerSer	176
Db	421	GCGAGGAGGTGTGGAGGAGGAGAGTGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG	480
QY	177	SerProLeuArgTyrAsnArgGlnIleGlyLysPheLeuValThrArgAlaGlyLeuTyr	196
Db	481	AGCCCTCGCCCTACAAACCCGAGATCGGGAGTTTATAGTACCCCGGCTGGCTCTAC	540
QY	197	TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu	216
Db	541	TACCTGACTGTGAGTGTCACCTTGTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG	600
QY	217	LeuValAspGlyValaLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSer	236
Db	601	CTGGTGAATGGTGTGTGGCTCGCTGCTGGAGGAAATTCAGCCACTGCGGCGAGT	660
QY	237	SerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGly	256
Db	661	TCCCTCGGCGCCAGCTCGGCTGTGCCAGGTGTCTGGGCTGTGGCCCTGCGGCGAGGG	720
QY	257	SerSerLeuArgIleArgTThrLeuProTrrPalaHisLeuLysAlaAlaProPheLeuThr	276
Db	721	TCCCTCGGCGAGTCCGAGCCGCTCCCTCGGAGGAAATTCAGCCACTGCGGCGAGT	780
QY	277	TyrPheGlyLeuPheGlnValHis	284
Db	781	TACTTCGAGCTCTTCAGGTTTAC	804
RESULT 3			
BC019047			
LOCUS	BC019047	1651 bp	mRNA linear PRI 11-DEC-2001
DEFINITION	Homo sapiens, similar to tumor necrosis factor (ligand) superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA, complete cds.		
ACCESSION	BC019047		
VERSION	BC019047.1 GI:17512138		
KEYWORDS	MGC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens. Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		

REFERENCE	1 (bases 1 to 1651)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	
COMMENT	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgabs-r@mail.nih.gov">cgabs-r@mail.nih.gov</a> Tissue Procurement: Louis Staudt CDNA Library Preparation: Rubih Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/nisc_mgcengr1.nih.gov">http://www.nisc.nih.gov/nisc_mgcengr1.nih.gov</a> Contact: Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., BenJamln, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stanlidrop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
FEATURES	
source	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAL Plate: 30 Row: P Column: 5 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596. Location/Qualifiers 1..1651 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:20669 IMAGE:4766071" /tissue_type="Primary B-cells from Tonsils" /clone_1lb="NIH_MGC_48" /lab_host="DH10B-R" /note="Vector: pOTB7" 106..510 /codon_start=1 /product="Similar to tumor necrosis factor (ligand) superfamily, member 12" /protein_id="AAH19047.1" /db_xref="GI:17512139" /translation="MAARSQRSGRRRPGPTALVPLALGLIALACILGLAVSLGSRASLSAODELVAEEDDPSELPMLPQTESSDPAFLNRLVRRRSAPKPKRTRARRAIATHEVHPREGQDAQNDG67TTCLRP"
CDS	
BASE COUNT	344 a 517 c 481 g 309 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3,53e-61
Score:	1331.50
Percent Similarity:	77.25%
Best Local Similarity:	77.25%
Query Match:	92.21%
DB:	9 Gaps: 1

US-09-245-198a-4 (1-284) x BC019047 (1-1651)			
QY	9	SerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaValArg	28
Db	25	TCCGCCCGCGGCGTCCCTCCCGCATCCCTCGGGATCCCGGGATGGGGCGGTGAGG	84
QY	29	GlnAlaGlnProProAlaProMetAlaAlaValArgArgSerGlnArgArgGlyArgArg	48
Db	85	CAGGCACAGCCCCCGCCCATGCGCGCCCTCGGAGAGCAGAGCGGAGGGCGCGCG	144
QY	49	GlyGluProGlyThrAlaLeuLeuValaProLeuAlaLeuGlyGlyGlyGlyAlaLeuAla	68
Db	145	GGGAGGCGGGGACCGCCCTGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	204

## REFERENCE

1 (bases 1 to 1651)  
 Strausberg, R.  
 Direct Submission  
 Submitted (07-DEC-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
 Tissue Procurement: Louis Staudt  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/nisc.mgcenhril.nih.gov>  
 Contact: [nisc.mgcenhril.nih.gov](mailto:nisc.mgcenhril.nih.gov)

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J.,  
 Tlonsong, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>  
 Series: IRAL Plate: 30 Row: P Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA g1: 4507596.

## FEATURES

## source

1..1651

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:20669 IMAGE:4766071"  
 /tissue\_type="Primary B-Cells from Tonsils"  
 /clone\_id="NIH\_MGC\_48"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 106..510  
 /codon\_start=1  
 /product="Similar to tumor necrosis factor (ligand)  
 superfamily, member 12"  
 /protein\_id="AAH19047.1"  
 /db\_xref="GI:17512139"  
 /translation="MAARRSQRRRGRGEPDTALVPLALGLIALACGLLAVSL  
 GSRASLSAOPAEIYAEEDDPSELPOTEEODPAPPLNRLVRRRSAPRGKTR  
 ARAIAAHYVHPRPDGDGAADGGYTCRLRP"

## CDS

BASE COUNT 344 a 517 c 481 g 309 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 3.53e-61 Length: 1651  
 Score: 1331.50 Matches: 275  
 Percent Similarity: 77.25% Conservative: 0  
 Best Local Similarity: 77.25% Mismatches: 1  
 Query Match: 92.21% Indels: 81  
 DB: 9 Gaps: 1

US-09-245-198a-4 (1-284) x BC019047 (1-1651)

QY	9	SerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaValArg	28
Db	25	TCCGCCCGCCGCTCCCTCCCGATCCCTCGGATCGCGGATGGGGGGCGGTGAGG	84
QY	29	GlnAlaGlnProProAlaProMetaAlaAlaArgArgSerGlnArgArgArgGlyArgArg	48
Db	85	CAGGCACAGCCCCCGCCCATGCGCCGCTGGAGACCAAGAGGAGGGGGCGCCGG	144
QY	49	GlyGluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAla	68
Db	145	GGGAGCGCGGCGCCCTGCTGCTCGCGCTCGCGCTGGGCTGGCGCTGGCC	204

Oy	69	CysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSerLeuSerAlaGln	88
Db	205	TGCCCCGGCCTCTCTCTGGCCGTGGTCAAGTTTGGGGAGCCGGGCGATCGCTGTCGCCCCAG	264
Oy	89	GIuProlAGlnGlnGluLeuValAlaGlnGluAspGlnAspProSerGluLeuAsnPro	108
Db	265	GAGCCTGCCAGAGAGAGCTGGTGGCAGAGAGACAGACACCCTCGGAACGTGAATCCC	324
Oy	109	GIuThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArgArg	128
Db	335	CAGACAGAAAGAACGAGATCTCTCGCTTTCCTGAACCCGACTGATTCGGCGTCGCAGA	384
Oy	129	SerAlaProLySGlyArgGlyThrArgAlaArgArgAlaIleAlaIleHisTyrGluVal	148
Db	365	AGTGCACCTTAAGAGCCCGGAACAGCGGCTCGAAGAGCATGGCAGCCCATTAATGAAGTT	444
Oy	149	HisProArgProGlyGlnAspGlyAlaGlnAla	159
Db	445	CATCCACGACCTGGACAGAGACGACCGCAGGC - AGATGGAGTTACACACTTGTCTAG	503
Oy	159		159
Db	504	GCCATGAGATACTAAGTGGTGGAGCCAAAGATTTGAACCCAGCTAGAATGTGCTGTAC	563
Oy	159		159
Db	564	TGGACATGTCATGATGAAGCGCAGGTGGCTGTGTCAGGGGTGAGGGTCCATCCAGGG	623
Oy	159		159
Db	624	GCCACATCCAAAAGGGGAGAGGAAGTTTCCAGAAAGAGAGAGACACATCTCCACC	683
Oy	160		160
Db	684	ATTACCAGAGGGCTCAAGAGAGAACAGACCAGGCTGGAGCGGACAGATGAGTGGCTGG	743
Oy	169	GIuGluAlaArgIleAsnSerSerSerProLeuArgTyrAsnArgGlnIleGlyGluPhe	188
Db	744	GAGGAGCCAGAGATACACAGCTCCAGCCCTCTGGCCTACACGCCACAGTCGGGAGATT	803
Oy	189	IleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluLys	208
Db	804	ATAGTCACCCGGGCGCTCTACTACCTGTACTGTCAAGTGCATCTTGATGAGGGGAAG	863
Oy	209	AlaValTyrLeuLeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGlu	228
Db	864	GCTGCTACCTGAACCTGACTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGG	923
Oy	229	GIuPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuGlyGlnValSer	248
Db	924	GAATTCTCAGCCACATCGCGCGAGTTCCTCGGCGCCCAAGCTCGGCTCTGCCAGGTGCT	983
Oy	249	GlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProThrAlaHis	268
Db	984	GGCGCTGTGGCCCTCGGCGCACAGGCTCTCCCTCGCGAGATCCGACCCCTCCCGGGCCCAT	1043
Oy	269	LeuLysAlaIleProPheLeuThrTyrPheGlyLeuPheGlnValHis	284
Db	1044	CTCAAGGCTGCCCTCTCTCACTTCAGACTTCGAGACTTTCAGGTTCCAC	1091
RESULT 4			
AF030099			
LOCUS	AF030099	1306 bp	mRNA
DEFINITION	Homo sapiens	TWEAK mRNA, complete cds.	linear
ACCESSION	AF030099		
VERSION	AF030099.1	GI:2707218	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 1306)		

AUTHORS	
TITLE	Chenoporphyrin, Bourdon, P., Xu, H., Hsu, Y.M., Scott, H., Hession, C., Garcia, I. and Browning, J.L.
JOURNAL	TWAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis
MEDLINE	J. Biol. Chem. 272 (51), 32401-32410 (1997)
PUBMED	98070415
	9405449
REFERENCE	2 (bases 1 to 1306)
AUTHORS	Bourdon, P., Hession, C., Tizard, R. and Browning, J.
TITLE	Direct Submission
JOURNAL	Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA
FEATURES	
source	Location/Qualifiers 1..1306 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17p13" /tissue_type="tonsil" 18..767 /note="Ligand in the TNF family; secreted protein; start codon not verified experimentally"
CDS	
BASE COUNT	247 a 434 c 368 g 257 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1.53e-59 Length: 1306
Score:	1298.00 Matches: 254
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatch: 0
Query Match:	89.89% Indels: 0
DB:	Gaps: 0
US-09-245-198A-4 (1-284) x AF030099 (1-1306)	
OY	31 GlnProProlAlaPrometAlaAlaArgSerGlnArgArgArgGlyGlu 50
Db	 3 CAGCCCCCGCCGCCATGGCCGCCTGGAGCAGAGAGCGAGGGCGCCGGGGAG 62
OY	51 ProGlyThrAlaLeuLeuValProleuAlaIleuGlyLeuGlyLeuAlaLeuCysLeu 70
Db	 63 CGGGCACCGCCGTGCTGTGTCGCCGCTGGGCTGGGCTTGCGCTGGCTTCCCTC 122
OY	71 GlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluPro 90
Db	 123 GGCCCTCGTGGCCCGTGTGCAGTTGGGGAGCCGGGCATGCTGCCGCCAGAGCCT 182
OY	91 AlaGlnGluGluLeuValAlaGluGlnAspGlnAspProSerGlyLeuAsnProGlnThr 110
Db	 183 GCCCAGGGAGGCTGTGTGCAGAGAGACAGACCCTGTGGAACTGAATCCCAGACA 242
OY	111 GluGlnSerGlnAspProlAlaProPheLeuAsnArgLeuValArgProArgArgSerAla 130
Db	 243 GAAGAAGCCAGCATCTCGGCCCTTTCTGAAACGCACTAGTTCGGCCTCGCAGAAATGCA 302
OY	131 ProLysGlyArgGlySerThrArgAlaArgArgAlaIleAlaIleAlaHisTryGluValHisPro 150
Db	 303 CCTAAAGSCGGAATAACACGCGGCTCGAAGAGCAGATCCGACCCCTATTATGAAGTTCATCCA 362
OY	151 ArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValAspGlyTyrTrpGluGlu 170
Db	 363 CGACCTGGACAGGAGCGGACGAGCAGTGTGACGGGACGGGACAGTAGTGTGCTGGAGGAA 422
OY	171 AlaArgIleAsnSerSerSerProLeuArgTrgYrAsnArgGlnIleGlyGluPheIleVal 190

```

|||||
Db 423 GCCAAGAACACAGCTCCAGCCCTCTGGCTACACACCGCCGAGATCGGGGATTTATATC 482
|||
Qy 191 ThrArgAlaGlyLeuYrYrLeuYrCysGlnValHisPheAspGlyLysAlaVal 210
|||
Db 483 ACCGGGCTGGGCTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 542
|||
Qy 211 TyrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuArgCysLeuGlnGluPhe 230
|||
Db 543 TACCTGAAGCTGGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
|||
Qy 231 SerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeu 250
|||
Db 603 TCAGCCACTGGGGCCAGTTCCTCGGGGCCAGCTCCGCTCCAGTCTGTGGGCTG 662
|||
Qy 251 LeuAlaLeuArgProGlySerSerLeuArgGlnLeuArgThrLeuProThrAlaHisLeuLys 270
|||
Db 663 TTGGCCCTGGCGCCAGGGTCTCCCTGGGATCCGACACCTCCCTGGGCCCATCTCAG 722
|||
Qy 271 AlaAlaProPheLeuThrYrPheGlyLeuPheGlnValHis 284
|||
Db 723 GCTGCCCTTCTCTCCTACTCTGACTTTCAGAGTTTAC 764
|||
RESULT 5
ARI40407 1236 bp DNA linear PAT 16-JUN-2001
LOCUS Sequence 1 from patent US 6207642.
DEFINITION ARI40407
ACCESSION ARI40407
VERSION ARI40407.1 GI:14482903
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Willey,S.R.
TITLE Member of the TNF family useful for treatment and diagnosis of
disease
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
source 1..1236
location/Qualifiers
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN
Alignment Scores:
Pred. No.: 7 51e-58 Length: 1236
Score: 1265.00 Matches: 248
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.60% Mismatches: 0
Query Match: 87.60% Indels: 0
DB: Gaps: 0
US-09-245-198a-4 (1-284) x ARI40407 (1-1236)
Qy 36 MetaAlaAlaArgArgSerGlnArgArgArgGlyArgGlyGluProGlyThrAlaLeu 55
|||
Db 1 ATGGCCGGCCGCTGGAGCAGAGAGGAGGGGGCGGGGGAGCCGGCAGCCGCCCTG 60
|||
Qy 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75
|||
Db 61 CTGTCCCGCGCGCGCTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 120
|||
Qy 76 ValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeu 95
|||
Db 121 GTGGTCACTTTGGGAGAGCGGGCATTCCTGTCCGCCAGGAGCTTGCCACGAGAGCTG 180
|||
Qy 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAsp 115
|||
Db 181 GTGGCAGAGGAGGAGCAGGAGCCGTGGGAATGATCCCGACAGAGAGAGAGAGAT 240
|||
Qy 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135
|||
Db 241 CCTGCGCTTCTCTGAAACGCACTAGTTCGGCTCGAAGAACTGCACCTAAAGCCGGA 300
|||

```

```

Qy 136 ThrArgAlaArgArgAlaAlaAlaHisGlyGluValHisProArgProGlyGlnAsp 155
|||
Db 301 ACAGGGCTCGAAGAGGATCGCAGCCATTATAGTTATCCAGCAGCTGGCAGAGG 360
|||
Qy 156 GlyAlaGlnAlaGlyValAspGlyThrValSerGlyYrPrpGluGlnAlaArgLysSer 175
|||
Db 361 GGAGCGCAGAGGTGTGTGAGCGGAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
|||
Qy 176 SerSerProLeuArgYrAsnArgGlnLeuGlyGluPheLeuValThrArgAlaGlyLeu 195
|||
Db 421 TCCAGCCCTGCGCCTCAACACCCAGATCGGAGATTATATGACACCGCGGCTG 480
|||
Qy 196 TyrYrLeuYrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAsp 215
|||
Db 481 TACTACCTGTACTCTCAGGTCTTGTATAGAGGAGGCTGTCTCTCACTGAACCTGAG 540
|||
Qy 216 LeuLeuValAspGlyValAlaLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAla 235
|||
Db 541 TTGGTGGTGGATGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
|||
Qy 236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 255
|||
Db 601 AGTCCCTCGGGCCCGCAGCTCCGCTCTGCGCAGGTGTCTGGGCTGTGGCCCTCGGCA 660
|||
Qy 256 GlySerSerLeuArgGlyLeuArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeu 275
|||
Db 661 GGTCTCTCTCTCGGAGATCCGACCTCCCTGGGCCCATCTCAAGCTGCCCTTCTCTC 720
|||
Qy 276 ThrYrPheGlyLeuPheGlnValHis 284
|||
Db 721 ACCTACTCGGAGCTCTCCAGTTTAC 747
|||
RESULT 6
AXI80714
LOCUS Sequence 1 from Patent WO0145730.
DEFINITION AXI80714
ACCESSION AXI80714
VERSION AXI80714.1 GI:15132570
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 898)
AUTHORS Willey,S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
FEATURES
source 1..898
location/Qualifiers
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN
Alignment Scores:
Pred. No.: 1.19e-47 Length: 898
Score: 1066.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.82% Indels: 0
DB: Gaps: 6

```

US-09-245-198a-4 (1-284) x AX180714 (1-898)

QY 78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuValAla 97  
 Db 250 AGTTGGGGAGCCGGGATGCTGTCCGCCAGAGAGCTGCCCGAGAGAGTGTGGCA 309  
 QY 98 GlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGlnGlnSerGlnAspProAla 117  
 Db 310 GAGGAGAGACAGACAGCCGCTGGAGATGAATCCCGACAGAAAGAACCCAGAGATCCCGC 369  
 QY 118 ProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThrArg 137  
 Db 370 CCTTCTGTAACCGACTAGTTCCGCCCTCGCAGAAGTCCACTTAAGGCCGGAAACACGG 429  
 QY 138 AlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAla 157  
 Db 430 GCTCGAAGACGATCCGACGCCCATTTATGAAGTTCAATCCAGACTGGACAGAGAGGCG 489  
 QY 158 GlnAlaGlyValAspGlyThrValSerGlyTyrGlnGlnGlnAlaArgIleAsnSerSer 177  
 Db 490 CAGGAGAGTGTGGACGGACAGTGAAGTGGCTGGAGAGAACCCAGAAACACAGCTCCAGC 549  
 QY 178 ProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyr 197  
 Db 550 CCTCGCGGTACAAACCGCCAGATCGGGAGTTTATAGTCAACCCGGGCTGCTCTACTAC 609  
 QY 198 LeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 217  
 Db 610 CTGTACTGTGACGTGACCTTTGATGAGGGAGAGGCTGTCTACTGAAGCTGAGCTGTG 669  
 QY 218 ValAspGlyValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaAlaSerSer 237  
 Db 670 GTGATGCTGTGTGCTGCCCTGCCCTGCTGGAGAAATCTCACCGCCAGCTGGCTCTCC 729  
 QY 238 LeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySer 257  
 Db 730 CTCGGGCCCCAGCTCCGCTCTCCAGGTGTCTGGGCTGTGGCCCTGGCGCCAGGGTCC 789  
 QY 258 SerLeuArgIleArgThrLeuProTTPAlaHisLeuLysAlaAlaProPheLeuThrTyr 277  
 Db 790 TCCCTCGGATCCGACACCTCCCTGGGCCATCTCAAGGCTGCCCTTCTCTCACCTAC 849  
 QY 278 PheGlyLeuPheGlnValHis 284  
 Db 850 TTGGACTCTTCACAGTTCAC 870

RESULT 7  
 AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
 LOCUS AF030100 Mus musculus TWEAK mRNA, partial cds.  
 DEFINITION AF030100  
 ACCESSION AF030100  
 VERSION AF030100.1 GI:2707220  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1168)  
 Chleporliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,  
 Hession,C., Garcia,I. and Browning,J.L.  
 TWEAK, a new secreted ligand in the tumor necrosis factor family  
 that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32401-32410 (1997)

JOURNAL  
 MEDLINE  
 PUBMED  
 98070415  
 9405449  
 2 (bases 1 to 1168)  
 Chleporliche,Y., Bixler,S., Tizard,R. and Browning,J.  
 Direct Submission  
 JOURNAL  
 Title  
 Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,  
 Cambridge, MA 02142, USA  
 FEATURES  
 Location/Qualifiers  
 1..1168

## CDS

/organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_type="peritoneal macrophages"  
 <1..679  
 /note="Ligand in the TNF family; secreted protein"  
 /codon\_start=2  
 /product="TWEAK"  
 /protein\_id="AAC53517.1"  
 /db\_xref="GI:2707221"  
 /translation="VLSGLALACGLLTVVSLGSMATLSAQPSELTAEADREP  
 PELNQTSESDVPEFLBOLVPRRSAPKRRARRRAIAHYEVHPPRQDQAGV  
 DCTVQMEFTKINSSPLRVDRIQIEFTYIRAGLYLYLCQVHFDGKAVLYLIDLIVN  
 GYLAIRCLDEFSATASSPQRLRCQVSGLLPLRPSSLRIRLPLMAHLKAPLTY  
 FGLFQVH"

BASE COUNT 242 a 360 c 298 g 268 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 3,79e-45 Length: 1168  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 70.64% Indels: 0  
 DB: Gaps: 0

US-09-245-198a-4 (1-284) x AF030100 (1-1168)

QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGly 80  
 Db 5 CTGAGCTGGGCGCTGGCGCTGGCTGCTCTGTGGCTCTGTGCTGTGCTGACGCTGGGG 64  
 QY 81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGlnGluAsp 100  
 Db 65 AGCTGGGCAACGCTGTCTCCAGAGCTTCTCAGAGAGAGCTGACAGCAGAGACGCC 124  
 QY 101 GlnAspProSerGlnLeuAsnProGlnThrGlnGlnGlnSerGlnAspProAlaProPheLeu 120  
 Db 125 CGGAGACCCCTGAACTGATCCCGACAGAGAGAAAGCAGAGTGTGTACTTCTTGG 184  
 QY 121 AsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArg 140  
 Db 185 GAACAACGTGTCGGGCTCGAAGAACTGCTCTAAAGCCGGAAGCGCGCTGGCGCA 244  
 QY 141 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
 Db 245 GCTATTTGCAAGCCCATTTATGAGGTTTCATCTCCGCCAGACAGAGATGAGCACACAGGT 304  
 QY 161 ValAspGlyThrValSerGlyTyrGlnGlnGlnAlaArgIleAsnSerSerProLeuArg 180  
 Db 305 GTGATGGACAGTGAAGTGGCTGGAGAGACCAAAATCAACAGCTCCAGCCCTTGGCG 364  
 QY 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
 Db 365 TACGACCGCCAGATTGGGAATTTACAGTCACAGGCTGGGCTCTACTACTGTACTGT 424  
 QY 201 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuLeuValAspGly 220  
 Db 425 CAGGTGCACTTTGATGAGGAAAGCTGTCTCTCTGAGAGCTGACTTCTGTGTAACGCT 484  
 QY 221 ValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaAlaSerLeuGlyPro 240  
 Db 485 GTGCTGGCCCTGCTGCTGGAGAAATTTCAAGCAGACAGCAAGCTCTCTGGGCC 544  
 QY 241 GlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 260  
 Db 545 CAGCTCCGTTTGTGCTGAGGTGTGGGCTGTGGCGCGCGGCGGCGGCTTCCCTTGG 604  
 QY 261 IleArgThrLeuProTTPAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 280  
 Db 605 ATCCGACCCCTCCCTGCGCTCATCTTAAGGCTGCCCTTCTTAACCTACTTGTGACTC 664  
 QY 281 PheGlnValHis 284  
 Db 1168



Db 665 TTTCAGTTCCAC 676

RESULT 8  
AC016876/c 177703 bp DNA linear HTG 06-AUG-2002

LOCUS AC016876.5 GI:21313830  
DEFINITION Homo sapiens clone RP11-186B7, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 10  
uncloned pieces.

ACCESSION AC016876  
VERSION AC016876  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Bouckgalter, B.,  
Brown, A., Castle, A., Colangelo, M., Collins, S., Collumore, A.,  
Cooke, P., Dearellano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M.,  
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,  
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Lenczky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N.,  
McMan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,  
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollard, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stenger-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
REFERENCE 3 (bases 1 to 177703)

AUTHORS Birren, B., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B.,  
Camarta, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Farso, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B.,  
Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrim, J., Meneses, L., Mihova, T., Mienga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Ruhankang, J., Roy, A., Schauer, S., Schnuppach, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Jun 2, 2002 this sequence version replaced gi:15421989.  
All repeats were identified using RepeatMasker:  
Smt, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIDR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
Project Information  
Center project name: L3849  
Center clone name: 186\_B\_7

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true origin of the pieces

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 48645: contig of 48645 bp in length
* 48646 48745: gap of 100 bp
* 48746 58277: contig of 9532 bp in length
* 58278 58377: gap of 100 bp
* 58378 65804: contig of 7427 bp in length
* 65805 65904: gap of 100 bp
* 65905 79793: contig of 13888 bp in length
* 79794 79893: gap of 100 bp
* 79894 99493: contig of 19600 bp in length
* 99494 99593: gap of 100 bp
* 99594 111049: contig of 11456 bp in length
* 111050 111149: gap of 100 bp
* 111150 125020: contig of 13871 bp in length
* 125021 125120: gap of 100 bp
* 125121 145109: contig of 19989 bp in length
* 145110 145209: gap of 100 bp
* 145210 169458: contig of 24249 bp in length
* 169459 169558: gap of 100 bp
* 169559 177703: contig of 8145 bp in length.
*
* Location/Qualifiers
*   1..177703
*     /organism="Homo sapiens"
*     /db_xref="taxon:9606"
*     /clone="RP11-186B"
*     /clone_11b="RP11-11 Human Male BAC"
42862 a 45967 c 46568 g 41254 t 1052 others
BASE COUNT
ORIGIN

```

Alignment Scores:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
1,398-20	589.50	65.62%	65.10%	40.82%	2
Length: 177703	Matches: 125	Conservative: 1	Mismatches: 0	Indels: 66	Gaps: 1
US-09-245-198A-4 (1-284) x AC016876 (1-177703)					
Oy 159	AlaGlyValAsp1ythrValSerGlyTrpGluGlnAlaArgIleAsnSerSerPro	178	178	178	178
Db 37636	TCAAGTGTGACGCGGACAGTGTGCTGGGAGGAGCCAGATCAACAGCTCCAGCCT	37577	37577	37577	37577
Oy 179	LeuArgTrpAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeu	198	198	198	198
Db 37576	CTGGGCTACACACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTCTACTACTG	37512	37512	37512	37512
Oy 199	TyrTys-----	200	200	200	200
Db 37516	TACGTGTAGGTAAAGCCCATCTGGCGTCAGTGCTAACCCACTAAGAGAGTGGGAAGGG	37455	37455	37455	37455
Oy 200	-----	200	200	200	200
Db 37456	TTGCCAGAGAGTGGGGACAAGCTACAGGCTGGAGGCTGAGTTGGGTTGGGTGG	37397	37397	37397	37397
Oy 200	-----	200	200	200	200
Db 37396	ATGGGATGCCCTGGCTGAGGAATTGGAATTGAGCGAGGCGAGAGGCTGG	37337	37337	37337	37337
Oy 201	-----GlnValHisPheSpgIuLysAlaValTyrLeuLys-----	213	213	213	213
Db 37336	ACTGGCCTGTGTCCCAACCCAGGTGACTGTATGAGGGAAGGCTGTACTAGAA	37277	37277	37277	37277
Oy 213	SleuAspLeuValAspGlyValLeuAlaLeuArgCysLeuGluIubPheSerAlaIth	233	233	233	233
Db 37276	GCTGGACTTCTGTGTGATGGTGTCTGGCCCTGGCTGGAGGAATTCACAGCAC	37217	37217	37217	37217
Oy 233	AlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLe	253	253	253	253

Db 37216 TGGGGAGATTCCCTGGGGCCAGCTCCGCTCTGCGCAGTGTGCTGTGGCCCT 37157  
QY 253 uatgproglyserSerleuargileargthrleuprotfralalhisleulysalaalapr 273  
Db 37156 GCGGCAAGGCTCTCCCTGGGATCGCACCCCTCCCTGGGCCATCTCAAGCTGCCCC 37097  
QY 273 ophleuThrTyPheGlyLeuPheGlnValHis 284  
Db 37096 CTTCCCTACTACTTCGACTCTTCCAGGTTAC 37063  
RESULT 9  
AC127470 215795 bp DNA linear HTG 17-JUL-2002  
LOCUS AC127470  
DEFINITION Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15  
unordered pieces.  
AC127470  
AC127470.1 GI:21886866  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS chimpanzee.  
SOURCE Pan troglodytes  
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
REFERENCE 1 (bases 1 to 215795)  
AUTHORS Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,  
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,  
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,  
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Lalic,P.,  
Lee-Iljin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B.,  
Marquillas,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,  
McCloskey,J.C., McDowell,J.J., Pegurigan,C., Pearson,R.,  
Portnoy,M.E., Prasad,A., Schueler,M.G., Stanfrilop,S., Thomas,J.W.,  
Thomas,P.J., Touchman,J.W., Tsurgueon,C., Vogt,J.L., Walker,M.A.,  
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 215795)  
TITLE Direct Submision  
REFERENCE Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717  
AUTHORS Grovmont Circle, Gailthersburg, MD 20877, USA  
JOURNAL ----- Genome Center  
COMMENT Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc.zoo@nih.gov  
----- Project Information  
Center project name: cms  
Center clone name: 145D13  
----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 203929 bases at least Q40  
Consensus quality: 206865 bases at least Q40  
Consensus quality: 209181 bases at least Q20  
Insert size: 190000; agarose-fp  
Quality coverage: 7.30x in Q20 bases; agarose-fp  
Quality coverage: 6.47x in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 3593: contig of 3593 bp in length  
\* 3594 3693: gap of unknown length  
\* 3694 8001: contig of 4308 bp in length  
\* 8002 8101: gap of unknown length

FEATURES  
source  
\* 8102 15639: contig of 7538 bp in length  
\* 15640 15739: gap of unknown length  
\* 15740 24775: contig of 9036 bp in length  
\* 24776 24875: gap of unknown length  
\* 24876 30150: contig of 5275 bp in length  
\* 30151 30250: gap of unknown length  
\* 30251 39212: contig of 8962 bp in length  
\* 39213 39312: gap of unknown length  
\* 39313 49342: contig of 10030 bp in length  
\* 49343 49443: gap of unknown length  
\* 49443 60956: contig of 11514 bp in length  
\* 60957 61056: gap of unknown length  
\* 61057 74520: contig of 13464 bp in length  
\* 74521 74620: gap of unknown length  
\* 74621 89522: contig of 14902 bp in length  
\* 89523 89622: gap of unknown length  
\* 89623 105537: contig of 15915 bp in length  
\* 105338 105637: gap of unknown length  
\* 105638 121197: contig of 15560 bp in length  
\* 121198 121297: gap of unknown length  
\* 121298 138883: contig of 17586 bp in length  
\* 138884 138983: gap of unknown length  
\* 138984 166802: contig of 27819 bp in length  
\* 166803 215795: gap of unknown length  
\* 166903 215795: contig of 48893 bp in length.  
Location/Qualifiers  
1. 215795  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-145D13"  
/clone\_1fb="RP43"  
1. 3593  
/note="assembly-fragment"  
misc\_feature 3694 . 8001  
/note="assembly-fragment"  
misc\_feature 8102 . 15639  
/note="assembly-fragment"  
misc\_feature 15740 . 24775  
/note="assembly-fragment"  
misc\_feature 24876 . 30150  
/note="assembly-fragment"  
misc\_feature 30251 . 39212  
/note="assembly-fragment"  
misc\_feature 39313 . 49342  
/note="assembly-fragment"  
misc\_feature 49443 . 60956  
/note="assembly-fragment"  
misc\_feature 61057 . 74520  
/note="assembly-fragment"  
misc\_feature 74621 . 89522  
/note="assembly-fragment"  
misc\_feature 89623 . 105537  
/note="assembly-fragment"  
misc\_feature 105638 . 121197  
/note="assembly-fragment"  
misc\_feature 121298 . 138883  
/note="assembly-fragment"  
misc\_feature 138984 . 166802  
/note="assembly-fragment"  
misc\_feature 166903 . 215795  
/note="assembly-fragment"  
BASE COUNT 52160 a 54859 c 53752 g 53513 t 1511 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 1 7e-20 Length: 215795  
Score: 589.50 Matches: 125  
Percent Similarity: 65.62% Conservative: 1  
Best Local Similarity: 65.10% Mismatches: 0  
Query Match: 40.82% Indels: 66  
DB: 2 Gaps: 1  
US-09-245-198a-4 (1-284) x AC127470 (1-215795)



[illegible]

db 113989 COAGTTCAC 113980

RESULT 11  
AC126925/c  
LOCUS AC126925 161428 bp DNA linear HTG 10-JUL-2002  
DEFINITION Canis familiaris clone Rp81-332E11, WORKING DRAFT SEQUENCE, 15  
unordered pieces.  
AC126925  
ACCESSION AC126925.1 GI:21724102  
VERSION HTG: HTGS\_PHASE1: HTGS\_DRAFT.  
KEYWORDS  
SOURCE  
ORGANISM Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
AUTHORS  
1 (bases 1 to 161428)  
Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,  
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,  
Haghighi, L., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,  
Lee, L.N.-O., Legaspi, R., Maduro, O.L., Maduro, V.B.,  
Margulies, E.H., Mastello, C., Maskeri, B., Mastrian, S.D.,  
McCluskey, J.C., McDowell, J., Paguirigan, C., Pearson, R.,  
Portnoy, M.E., Prasad, A., Schuler, M.G., Stantlippo, S., Thomas, J.W.,  
Thomas, P.J., Touchman, J.W., Tsurugeno, C., Vogt, J.L., Walker, M.A.,  
Weheby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 161428)  
Green, E.D.  
Direct Submission  
Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717  
Grovmont Circle, Gaithersburg, MD 20877, USA  
-----  
Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: [nisc-zoehghrl.nih.gov](mailto:nisc-zoehghrl.nih.gov)  
-----  
Project Information  
Center project name: cwp  
Center clone name: 332E11  
-----  
Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 148712 bases at least Q40  
Consensus quality: 151471 bases at least Q30  
Consensus quality: 153125 bases at least Q20  
Insert size: 152000; agarose-fp  
Insert size: 160028; sum-of-contigs  
Quality coverage: 5.36x in Q20 bases; agarose-fp  
Quality coverage: 5.10x in Q20 bases; sum-of-contigs  
-----  
NOTE: This is a 'working draft' sequence. It currently  
consists of 15 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.  
1 4734: contig of 4734 bp in length  
4735 4834: gap of unknown length  
4835 7325: contig of 2491 bp in length  
7326 7425: gap of unknown length  
7426 12261: contig of 4836 bp in length  
12262 12361: gap of unknown length  
12362 19031: contig of 6670 bp in length  
19032 19131: gap of unknown length  
19132 27849: contig of 8718 bp in length  
27850 27950: gap of unknown length  
27950 31801: contig of 3852 bp in length  
31802 31901: gap of unknown length  
31902 40781: contig of 8880 bp in length

```
* 40782 40881: gap of unknown length
* 40882 49783: contig of 8902 bp in length
* 49784 49883: gap of unknown length
* 49884 59150: contig of 9267 bp in length
* 59151 59250: gap of unknown length
* 59251 68679: contig of 9429 bp in length
* 68680 68779: gap of unknown length
* 68780 78984: contig of 10205 bp in length
* 78985 79084: gap of unknown length
* 79085 93446: contig of 14362 bp in length
* 93447 93546: gap of unknown length
* 93547 102175: contig of 8629 bp in length
* 102176 102275: gap of unknown length
* 102276 131685: contig of 29410 bp in length
* 131686 131785: gap of unknown length
* 131786 161428: contig of 29643 bp in length.
FEATURES
Source
1..161428
/organism="Canis familiaris"
/db_xref="taxon:9615"
/clone="RP81-332E11"
/clone_id="RP81"
1..4734
misc_feature
/note="assembly_fragment"
4835..7325
/note="assembly_fragment"
7426..12261
/note="assembly_fragment"
12362..19031
/note="assembly_fragment"
19132..27849
/note="assembly_fragment"
27950..31801
/note="assembly_fragment"
31902..40781
/note="assembly_fragment"
40882..49783
/note="assembly_fragment"
49884..59150
/note="assembly_fragment"
59251..68679
/note="assembly_fragment"
68780..78984
/note="assembly_fragment"
79085..93446
/note="assembly_fragment"
93547..102175
/note="assembly_fragment"
102276..131685
vector_side:right"
misc_feature
/note="assembly_fragment"
131786..161428
/note="assembly_fragment"
BASE COUNT 39586 a 41863 c 41233 g 37340 t 1406 others
ORIGIN
Alignment Scores:
Pred. No.: 7.39e-19 Length: 161428
Score: 555.50 Matches: 118
Percent Similarity: 64.92% Conservative: 6
Best Local Similarity: 61.78% Mismatches: 9
Query Match: 38.47% Indels: 58
DB: 2 Gaps: 1
US-09-245-198a-4 (1-284) x AC126925 (1-161428)
Oy 152 ProglYGLNAspGlyAlaGLAlaGlyValaAspGlyThrValSerGlyTrpGluGlnAla 171
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 42806 CCAATCTCTGAACCCCTGATTCAGGTGTGCGACGCGGAGTGTGGGAGAGAGGCC 42747
Oy 172 ArgIleAsnSerSerProLeuArgTyrAsnArgIleGlyGluPheIleValThr 191
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
DB 42746 AAATCAACAGCTCCACCCACTGCGCTATGACCCGACAGCGGGGAATTATAGTCAGC 42687
Oy 192 ArgAlaGlyLeuTyrTyrLeuTyrCys----- 200
DB 42686 CGGGCTGGCTCTACTACTGTACTGTACGACAGTAAGCCCCACCTGCTCCACGTAAGGCA 42627
Oy 200 ----- 200
DB 42626 GAACGGGGGGCCCCAGGTTGAGGACAGACGAGAGGGTTTGCTTCAGGGGGCAGTTGGG 42567
Oy 200 ----- 200
DB 42566 GTGGGAGCGTAGGTCGGGGCCCGATGGATCTGGGGTCACCAAGACCCCGAGCT 42507
Oy 201 -----GlnValHisPheAspGluGlyAlaValTyrLeuLysLe 214
DB 42506 GCGGCTCTCCGCGGCCCCCAGGTCTGATATAGGGAGGCTGCTACTCTCAAGCT 42447
Oy 214 uAspLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAl 234
DB 42446 GCACTTGTGTGTGATGACGCCCTGGCCCTGCTGCTGGAGAGTTCTCCGCCACAGC 42387
Oy 234 aAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuAr 254
DB 42386 CGCCAGCAGCGCTGGGCCCCCAGCTCCGCTCGCCAAAGTGTGGGCTTGCCCTCCG 42327
Oy 254 gProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPh 274
DB 42326 GCCCGGTCCTCCCGCGGATCCGACCCCTCCCGGCCATCTCAAGCGGCCCCCTT 42267
Oy 274 eleuThrTyrPheGlyLeuPheGlnValHis 284
DB 42266 CCTTACTACTTTCGACTCTTCCAGGTTAAC 42236
RESULT 12
AC069459/c 203083 bp DNA linear HTG 27-JUN-2001
LOCUS Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT
DEFINITION SEQUENCE, 7 unordered pieces.
ACCESSION AC069459 GI:14547768
VERSION AC069459.23
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 203083)
Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,
Dederich,D., Thomas,S., Okwuonu,G., Carlock,C., Garner,T.,
Addison,S., Pace,A., Williams,G., Bonnih,D., Brooks,A., Brown,J.,
Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,
Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,
Gorrell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hoques,M.,
Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,
Kovar,C., Liu,X., Liu,W., Louisgeed,H., Lozano,R.J., Martin,R.,
Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morris,S.,
Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Ogun,M., Parlish,B.,
Perez,L., Reltter,D., Say,J., Shen,H., Vasquez,L., Watlington,S.,
Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,
Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstock,G.,
Worley,K. and Gibbs,R.
Direct Submission
Unpublished
JOURNAL
TITLE
JOURNAL
AUTHORS
Worley,K.C.
REFERENCE
2 (bases 1 to 203083)
Direct Submission
Submitted (31-MAY-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced gi:12621364.
----- Genome Center
Center: Baylor College of Medicine
```

```

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: MAF0
Center clone name: RP23-168P5
-----
Summary Statistics
Sequencing vector: M13; L088821
Chemistry: Dye-terminator Big Dye 52% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 212648 bases at least Q40
Consensus quality: 218902 bases at least Q30
Consensus quality: 222384 bases at least Q20
Estimated insert size: 210656; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agatose-fp estimation
Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drafter.data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 62152: contig of 62152 bp in length
* 62153 62252: gap of unknown length
* 62253 118772: contig of 56520 bp in length
* 118773 118872: gap of unknown length
* 118873 148924: contig of 30052 bp in length
* 148925 149024: gap of unknown length
* 149025 167231: contig of 18207 bp in length
* 167232 167331: gap of unknown length
* 167332 169807: contig of 22576 bp in length
* 169808 190007: gap of unknown length
* 190008 196537: contig of 6530 bp in length
* 196538 196637: gap of unknown length
* 196638 203083: contig of 6446 bp in length.
*
* Location/Qualifiers
*     1..203083
*         /organism="Mus musculus"
*         /db_xref="taxon:10090"
*         /chromosome="11"
*         /clone="RP23-168P5"
*
BASE COUNT      52662 a 49293 c 47892 g 52633 t      603 others
ORIGIN
-----
Alignment Scores:
Pred. No.:      9,31e-19      Length:      203083
Score:          555.50        Matches:      117
Percent Similarity: 67.22%    Conservative: 4
Best Local Similarity: 65.00%  Mismatches:  5
Query Match:     38.47%       Indels:       54
DB:              2           Gaps:         1
-----
US-09-245-198A-4 (1-284) x AC069459 (1-203083)
OY      159 AlAglyValAspGlyThrValSerGlyTrpGluGluAlaArgIleAsnSerSerPro 178
Db 43305 TCAGGTGTGGATGGGACACTGATGGCTGGGAAGACCAAAATCAACAGCTCCAGCCCT 43246
OY      179 LeuAtgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTrpLeu 198
Db 43245 CTGGCGCTACGACCGCCAGATTGGGGAAATTACAGTCATCAGGCGTGGGCTTACTACTCG 43186
OY      199 TTTCTT
Db 43185 TACTCTCAAGTAAAGCCCTGCCCTGCTTCAAGGGGACAGCAAGGCTAAGGGAGAGAGCA 43126
OY      200 -----

```

```

Db 43125 CTGGCAGAGGAAATGCGTGGGAGGAGGAGACCTCGGACATCAGGAAGAGATGCGTGGCTT 43066
Oy 201 -----GlnValHisPheAs 205
Db 43065 TCGATGAGGAGGAGGCGAGAGGTGTGATTTGCTTGGCTCTCTCTGTTCCAGGTGCACATTTTGA 43006
Oy 205 pGluValLysAlaValIYrLeuLysLysAspLeuLeuValAspGlyValLLeuAlaLeuAr 225
Db 43005 TGAGGGAAGAAGGCTGTCTACCTGAAAGCTTGACTTGTGTGTGAACGGTGTGCTGGCCCTGCG 42946
Oy 225 gCysLeuLugLupheSerAlaThrAlaAlaSerSerLeuLylProGlnLeuArgLeuCy 245
Db 42945 CTGGCTGGAAGAATCTCAGGCACAGCAGCAAGCTCTCTGGGCCCCAGCTCCGTTTGTG 42886
Oy 245 sGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleAagThrLeuPr 265
Db 42885 CCAGGTGTCTGGGGCTGTTGGCCGCTGCGGCACAGGCTTTCCTTGGGATCCGACCCCTGCC 42826
Oy 265 oTPrpHisLeuLysAlaAlaProPheLeuThyrPheGlyLeuPheGlnValHis 284
Db 42825 CTGGGCTCATCTTAAGGTGCGCCCTTCTCAACCTAGCTTGAGACTTTTCAAGTTTCAAC 42768
RESULT 13
AL603707/c AL603707 234182 bp DNA linear ROD 17-NOV-2001
LOCUS Mouse DNA sequence from clone RP23-422L16 on chromosome 11,
DEFINITION complete sequence.
ACCESSION AL603707
VERSION AL603707.5 GI:17017790
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 234182)
Pearce/A.
Direct Submission
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced gi:16605765.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction enzyme digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; SW:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is
from the RPc1-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTORS: pBAC3.6
This sequence is the entire insert of clone RP23-422L16.
FEATURES
Source
Location/Qualifiers
1..234182
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_1lp="RPc1-23"
/complement(84050..84131)
misc.feature

```



misc\_feature /note="assembly\_fragment"  
8536..15799  
/note="assembly\_fragment"  
misc\_feature 15900..25224  
/note="assembly\_fragment"  
misc\_feature 25325..32504  
/note="assembly\_fragment"  
misc\_feature 32605..40970  
/note="assembly\_fragment"  
misc\_feature 41071..56590  
/note="assembly\_fragment"  
misc\_feature 56691..73769  
/note="assembly\_fragment"  
misc\_feature 73870..90859  
/note="assembly\_fragment"  
misc\_feature 90960..111428  
/note="assembly\_fragment"  
misc\_feature 111529..153553  
/note="assembly\_fragment"  
clone\_end:sp6  
vector\_side:right"

BASE COUNT 39178 a 37393 c 36616 g 39259 t 1107 others  
ORIGIN

Alignment Scores:  
Pred. No.: 9, 77e-16 Length: 153553  
Score: 495.00 Matches: 110  
Percent Similarity: 58.598 Conservative: 6  
Best Local Similarity: 55.568 Mismatches: 17  
Query Match: 34,288 Indels: 66  
DB: 2 Gaps: 1

US-09-245-198a-4 (1-284) x AC126921 (1-153553)

QY 152 ProGlylnaspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTTPGluAla 171  
||| .....  
Db 26848 CCGTCTCTAACCACCCACCATTCAGGTGTGAGGAGGAGGTGAGTGGAGAGGAGCC 26907  
||| .....  
QY 172 ArgIleasnSerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThr 191  
:::|||||.....  
Db 26908 AATATCAACAGCTCCAAACCCCTGGCTATGACGCCAGACCGGGCAATTACGGTACC 26967  
||| .....  
QY 192 ArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal..... 202  
||| .....  
Db 26968 CGGGCTGGGCTGTACTGTACTGTACAGT-AAGCCCACTGGCTCCACGGGTAAG 27026  
||| .....  
QY 202 ..... 202  
Db 27027 CCGGAAGTGAAGCAGAAAGCTGGGCTTGGGGGCGCAAGTTAAAGTGGGAGCG 27086  
||| .....  
QY 202 ..... 202  
Db 27087 GAGCGTGGGTTGGGCTGAGAGAGGAGCTTGGGCTCTAAGACACCTGAGATGAAGCC 27146  
||| .....  
QY 203 .....HisPheAspGlu 206  
Db 27147 AGGGCCAGCAGAGGCTGAGCTCGCCCTCCCTGCCGCCAGAGTGTGATGAG 27206  
||| .....  
QY 207 GilylValAlaValTyrLeuLysLeuAspLeuLeuValAspGlyValLeuAlaLeuArgCys 226  
||| .....  
Db 27207 GGGAAAGGCTCTTACTGAAAGCTGAGTCTGGTGGATATACACGCTGGGCTGGCTGC 27266  
||| .....  
QY 227 LeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGln 246  
||| .....  
Db 27267 CTGGAGGAATTCCTGGGCACTGGGCGCAGTCCCTGGGCGCCAGCTGCTGTGCAA 27326  
||| .....  
QY 247 ValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTyr 266  
||| .....  
Db 27327 GTGTCAAGGCTCTTGGCTGGGCGAGGGTCTTCCCTGGGAGATCCGACCTGCCAG 27386  
||| .....  
QY 267 AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284  
||| .....  
Db 27387 ACCCACTCAAGGCTGGCCCTTCCTACCTACTTGGAGCTTTCACAAGTTTAC 27440

RESULT 15  
AC119115  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-320N23, \*\*\* SEQUENCING IN PROGRESS  
\*\*\*, 32 unordered pieces.

AC119115  
AC119115.2 GI:21746718  
VERSION  
KEYWORDS  
HTG: HTGS\_PHSBL  
SOURCE  
Norway rat.  
Rattus norvegicus  
ORGANISM

REFERENCE  
AUTHORS  
Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbarta,J., Benton,J., Bimaga,K., Blankenburg,K., Bonin,D.,  
Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,R.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyte,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frazz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsso,E., Kelly,S., Khan,U., King,L., Korvah,J., Kover,C.,  
Kratovic,J., Kureshi,A., Landry,N., Deal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,M., Louised,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,S.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogund,M., Okwomu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojka,A., Rojokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,  
Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G., and Gibbs,R.

TITLE  
JOURNAL  
2 (bases 1 to 138792)  
Unpublished

REFERENCE  
AUTHORS  
Muzny K.C.

TITLE  
JOURNAL  
Submitted (25-APR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
Muzny K.C.

TITLE  
JOURNAL  
Direct Submission  
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

COMMENT  
On Jul 14, 2002 this sequence version replaced gi:20303440.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM



Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a>	
Contact: <a href="mailto:hgsc-help@bcm.tmc.edu">hgsc-help@bcm.tmc.edu</a>	
Project Information	
Center Project name: GUXG	
Center clone name: CH230-320N23	
Summary Statistics	
Sequencing vector: Plasmid	
Chemistry: Dye-terminator Big Dye: 100% of reads	
Assembly program: Phrap; version 0.990329	
Consensus quality: 112720 bases at least Q40	
Consensus quality: 116666 bases at least Q30	
Consensus quality: 119165 bases at least Q20	
NOTE: Estimated insert size may differ from sequence length.	
(see <a href="http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html">http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html</a> )	
NOTE: This is a 'working draft' sequence. It currently	
consists of 32 contigs. The true order of the pieces	
is not known and their order in this sequence record is	
arbitrary. Gaps between the contigs are represented as	
runs of N, but the exact sizes of the gaps are unknown.	
This record will be updated with the finished sequence	
as soon as it is available and the accession number will	
be preserved.	
1	1022: contig of 1022 bp in length
1023	1122: gap of unknown length
1123	2394: contig of 1272 bp in length
2395	2494: gap of unknown length
2495	3663: contig of 1169 bp in length
3664	3763: gap of unknown length
3764	4879: contig of 1116 bp in length
4880	4979: gap of unknown length
4980	6787: contig of 1808 bp in length
6788	6887: gap of unknown length
6888	8192: contig of 1305 bp in length
8193	8292: gap of unknown length
8293	9461: contig of 1169 bp in length
9462	9561: gap of unknown length
9562	11435: contig of 1874 bp in length
11436	11535: gap of unknown length
11536	13535: contig of 2000 bp in length
13536	13635: gap of unknown length
13636	17323: contig of 3688 bp in length
17324	17423: gap of unknown length
17424	20405: contig of 2983 bp in length
20407	20506: gap of unknown length
20507	23445: contig of 2939 bp in length
23446	23545: gap of unknown length
23546	25600: contig of 2055 bp in length
25601	25700: gap of unknown length
25701	28049: contig of 2349 bp in length
28050	28149: gap of unknown length
28150	30474: contig of 2325 bp in length
30475	30574: gap of unknown length
30575	35498: contig of 4924 bp in length
35499	35598: gap of unknown length
35599	39284: contig of 3686 bp in length
39285	39384: gap of unknown length
39385	41970: contig of 2586 bp in length
41971	42070: gap of unknown length
42071	45659: contig of 3589 bp in length
45660	45759: gap of unknown length
45760	50332: contig of 4573 bp in length
50333	50432: gap of unknown length
50433	54695: contig of 4263 bp in length
54696	54795: gap of unknown length
54796	58289: contig of 3494 bp in length
58290	58389: gap of unknown length
58390	63599: contig of 5210 bp in length
63600	63699: gap of unknown length
63700	70305: contig of 6606 bp in length
70306	70405: gap of unknown length
70406	76123: contig of 5718 bp in length
76124	76223: gap of unknown length

```

*      76224      84961:  confg of 8738 bp in length
*      84962      85061:  gap of unknown length
*      85062      93614:  confg of 8553 bp in length
*      93615      93714:  gap of unknown length
*      93715      103352:  confg of 9638 bp in length
*      103353      103452:  gap of unknown length
*      103453      110299:  confg of 6847 bp in length
*      110300      110399:  gap of unknown length
*      110400      116180:  confg of 5781 bp in length
*      116181      116280:  gap of unknown length
*      116281      125484:  confg of 9204 bp in length
*      125485      125584:  gap of unknown length
*      125585      138792:  confg of 13208 bp in length.
FEATURES
    source
        1. .138792
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
            /clone="CH230-320N23"
BASE COUNT      34391 a      31996 c      31712 g      36193 t      4500 others
ORIGIN
Alignment Scores:
Pred. NO.:      1.7e-15      Length:      138792
Score:      489.50      Matches:      102
Percent Similarity:      64.50%      Conservative:      7
Best Local Similarity:      60.36%      Mismatches:      7
Query Match:      33.90%      Indels:      53
DB:      2      Gaps:      1
US-09-245-198A-4 (1-284) x AC119115 (1-138792)
QY      160      G1YVALASPGLYTHVAISERGLYTRPGIUNLALARGLLEAANSERSESRPROLEU 179
DB 137345      GGTGTGGATGGAGACAGTAGGGCTGGGAAAGAACCAAAATCAACACTCCAGCCCTCTG 137404
QY      180      ARGTYRANATRGGINLIEGLYGLIPHEILVALHRAFGALAGLYLEUTRYTYRLEUTYR 199
DB 137405      CGCTATGACCCGCCAGATTGGGGAATTACGGTCACTACAGGGCTGGGCTCTACTACTCTAC 137466
QY      200      Cys----- 200
DB 137465      TGTCAAGTAAAGCCTGGGCTCCATGGTAGAGCATGGCTPAAAGGAGAAAGCCTTGGA 137524
QY      200      ----- 200
DB 137525      AGGAATGGGTGGAGGTGGAGAACCTGGGTTTCATGAAGAAGATGCGTGGATTTCGATG 137584
QY      201      ----- -G1YVALHISPASPGLU 206
DB 137585      AGGCAGGCAGAGGCTCTGGATTTCGTGCTTCTCTGCTCCAGGTCGACATTGGATGAG 137644
QY      207      G1YLSA1AVALTYRLEULYSLEUASPHEULEUVALASPGLYLALEUALALEUARGCYS 226
DB 137645      GGGAGAGCAAGTCTACTCAACTGAGACTGCTGGTGAATGATGTGTGTGGCCCTCGCTGC 137704
QY      227      LEUGLUGLUPHESERALATHALALASERLEUGLYPROG1NLEUARGLEUCYSGLN 246
DB 137705      CTGGAGAAGATTCTCAGCCACAGCAGCGAGCTCTCTGGGCCCCCAGCTCCGTCCTGTGCAG 137764
QY      247      VALSERGLYLEULEUALALEUARGPROGLYSERLEUARGILEATGYTHLEUPROTRP 266
DB 137765      GTGTGTGGGCTGTGGCTCTCTCGACCAAGGCTTCCTTCGGATCCGATACCATCTTGC 137824
QY      267      ALAHISLEULYSALAA1APROPHLEU 275
DB 137825      AGTCATATTAAAGCGGCTACTTCGTT 137851

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 00:47:43 ; Search time 294.043 Seconds  
(without alignments)  
2175.080 Million cell updates/sec

Title: US-09-245-198A-4  
Perfect score: 1444  
Sequence: 1 MSLLDFEISARRLPLPRSLG.....PMAHLKAAPLTYGFLGVH 284

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODL=frame+pn.model -DEV=xlh  
-Q=Cgn2\_1/USPTO/US09245198/runat\_24032003\_163552\_25944/app\_query.fasta.1.846  
-DB=N\_Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum2 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=200000000  
-USER=US09245198.ecgn.1.1-338\_etunat\_24032003\_163552\_25944 -NCPU=6 -ICPU=3  
-NO\_XLPPY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq.101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1444	100.0	1373	19	AAV18600	Homo sapiens tumou
2	1406	97.4	1364	24	ABK34881	Homo cDNA encodin
3	1377	95.4	1421	20	AAV56000	Human tumour necro
4	1366	94.6	1353	21	AAA49717	Human PRO207 cDNA
5	1366	94.6	1353	24	ABK40255	Human encoding huma
6	1265	87.6	1236	19	AAV47613	TNF related endothe
7	1265	87.6	1236	20	AAV47613	Human TREPA (TNF r
8	1066	73.8	898	22	AAV03964	Expression vector
9	1062	73.5	1030	24	AAV32424	Human TNFRL3 DNA.
10	1020	70.6	1168	19	AAV18599	Mus musculus tumou
11	968	67.0	701	20	AAV23425	Mouse TNFRL3 DNA.
12	127	8.8	2856	19	AAV24140	Homo sapiens BARD1
13	126.5	8.6	1660	21	AAA49205	Corn putative lecl
14	124.5	8.6	2832	20	AAV69289	Murine JNK-Interac
15	124.5	8.6	154746	22	AAV25519	Human herpesvirus
16	124	8.6	3957	22	AAV09686	HSV-2 immediate ea
17	124	8.6	154746	22	AAV25519	Human herpesvirus
18	123.5	8.6	11705	19	AAV62160	HSV-2 strain SB5 C
19	123.5	8.3	117213	19	AAV62176	CDNA of a human di
20	120.5	8.3	2828	24	AAA46674	Human benign prost
21	120.5	8.3	2828	24	ABK64503	Colon adenocarcino
22	120.5	8.3	2828	24	ABL61804	Thyroid cancer rel
23	117.5	8.1	3880	24	ABL67211	Human TATA-binding
24	117.5	8.1	3901	17	AAV42224	TATA-binding prote
25	117.5	8.1	3901	18	AAV79598	TATA-binding prote
26	117.5	8.1	3902	15	AAV07338	DNA encoding an ac
27	117	8.1	1030	20	AAV98230	Nucleotide sequenc
28	117	8.1	1030	21	AAV12514	Drosophila melanog
29	116	8.0	978	23	ABL21473	DNA encoding tumou
30	116	8.0	2148	24	ABK11680	Chicken beta-actin
31	115	8.0	1275	11	AAV01620	Expression vector
32	115	8.0	1345	12	AAV013318	Hybrid promoter of
33	115	8.0	1495	22	AAV030652	Chicken beta-actin
34	115	8.0	1495	22	AAV64455	Mouse Meg1/Grb10 n
35	115	8.0	4211	22	AAH21793	Rat expression pro
36	115	8.0	5504	24	AAV43172	Human expression p
37	115	8.0	5565	24	AAV43171	Human gene containi
38	115	8.0	5581	22	AAV41035	Positive selection
39	115	8.0	5759	24	ABK49521	Positive selection
40	115	8.0	6148	24	ABK49520	PCAGGS-HSEPOR2 rat
41	115	8.0	6256	24	AAV43173	Human p531-alpha-
42	115	7.9	8033	22	AAV12739	Human p531-alpha-
43	114.5	7.9	12700	19	AAV62133	HSV-2 strain SB5 C
44	113.5	7.9	3836	15	AAV05997	Vitamin D hydroxyl
45	113.5	7.9	44377	18	AAV78508	Platenolide syntha

## ALIGNMENTS

RESULT 1  
ID AAV18600 standard; cDNA: 1373 BP.

AAV18600:  
21-JUL-1998 (first entry)

DE Homo sapiens tumour necrosis factor related ligand (TRELL) gene.

KW TRELL: tumour necrosis factor related ligand; tnfr; treatment;

KW cancer; autoimmune disease; immune system; stimulation; suppression;

XX graft rejection; ds.

OS Homo sapiens.  
XX Key Location/Qualifiers  
FH CDS 1..852

```
FT      /*tag= a
XX      /note= "tumour necrosis factor related ligand"
PN      MO9805783-A1.
XX      12-FEB-1998.
PD      07-AUG-1997; 97WO-US13945.
XX      18-MAR-1997; 97US-0040820.
PR      07-AUG-1996; 96US-0023541.
PR      18-OCT-1996; 96US-0028515.
XX      (BIOJ ) BIOGEN INC.
PA      (UYGE-) UNIV GENEVA FACULTY MEDICINE.
PI      Browning JL, Chicheportiche Y;
XX      WPI; 1998-145619/13.
DR      P-PSDB; AAW47525.
XX      Tumour necrosis factor related ligand - useful for, e.g. treating
PT      cancer, auto-immune disease and immune responses to tissue grafts
XX      Claim 2; Pages 48-50; 69pp; English.
XX      The sequence is that encoding human tumour necrosis factor related
CC      ligand (TRELL). TRELL or active fragments can be included with a
CC      carrier in pharmaceutical compositions to treat cancer, autoimmune
CC      diseases or immune responses to tissue grafts, or to stimulate or
CC      suppress the immune system. It is useful to screen for TRELL
CC      receptors, by labelling with a detectable label and screening
CC      compositions for binding. Agents interfering with TRELL-receptor
CC      binding can also be screened for, can then be administered,
CC      optionally with interferon- gamma, to induce cell death or
CC      treat, suppress or alter immune responses (especially involving human
CC      adenocarcinoma cells) involving a signal pathway between TRELL and its
CC      receptor. The DNA sequence can be used in gene therapy for
CC      TRELL-related disorders in mammals (especially humans), e.g. tumours,
CC      autoimmune and inflammatory diseases or inherited genetic disorders,
CC      by introducing into cells, and expressing, therapeutically effective
CC      amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
CC      It may also be of use in the preparation of prepare probes for
CC      screening natural/synthetic DNAs for TRELL-encoding sequences
CC      and for antisense therapy.
XX      SQ      Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other:
XX
XX      Alignment Scores:
XX      Pred. No.:      9,06e-96      Length:      1373
XX      Score:      1444.00      Matches:      284
XX      Percent Similarity:      100.00%      Conservative:      0
XX      Best Local Similarity:      100.00%      Mismatches:      0
XX      Query Match:      100.00%      Indels:      0
XX      DB:      19      Gaps:      0
XX
XX      US-09-245-198a-4 (1-284) x AAV18600 (1-1373)
QY      1 MetSerLeuLeuAspPheGluIleSerAlaArgArgLeuProLeuProArgSerLeuGly 20
DB      1 ARGTCATTGTTAACATTTTGAATAATTCGCCGCCGCTCCCTCCCGCATCCCTCGG 60
QY      21 SerArgAspGlyValAlaValArgGlnAlaGlnProProAlaProMetValAlaArgArg 40
DB      61 TCCCGGGATGGGGGGCGGTGAAGCAGACGCCGCCGCCCAATGGCGCCCTCGG 120
QY      41 SerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeuValProLeuAla 60
DB      121 AGCCAGAGCGGAGGGGGGGCGCGGGGAGCCGGCAGCCGCTGCTGCTGCCGCTCGGC 180
QY      61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValAlaSerLeuGly 80
DB      181 CTGGGCTTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
```

```
QY      81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGluGluAsp 100
DB      241 AGCCGGGCAATCGCTGTCCGCCAGAGAGCTGCCAGAGAGAGCTGTGGCAGAGAGAGAC 300
QY      101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeu 120
DB      301 CAGGACCCGCTCGCACTGTAATCCCGAGACAGAAAGGAGAGATCTCGCTTTCTCG 360
QY      121 AsnArgLeuValArgProArgSerAlaProLysGlyArgLysThrArgAlaArgArg 140
DB      361 AACCGACTAGTTGCGCTTGCAGAACTGCACCTAAAGGCCGGAACACAGCGCTCGAAGA 420
QY      141 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160
DB      421 GCGATCGCAGCCCATTTATGAAGTTTCATCCAGACACGACGAGAGAGAGCGAGAGGT 480
QY      161 ValAspGlyThrValSerGlyTrrGluGluAlaArgIleAsnSerSerSerProLeuArg 180
DB      481 GTGGAGCGGACACTGAGTGGCTGGGAGAGAGCCAGATCAGATCCAGCCCTCTGCGC 540
QY      181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200
DB      541 TACACCGCCAGATCGGGAGTTTATAGTCACCCGGGCTGGCTTACTTACTTCTGT 600
QY      201 GlnValHisPheAspGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220
DB      601 CAGGTCACTTGTATGAGGGAGAGGCTGTCTACTGAGCTGAGCTTGTGATGTGT 660
QY      221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerLeuGlyPro 240
DB      661 GTCTGTGCGCTGGCGCTCGCGAGAAATTCAGCCACGCGGCGCATGCTCCGCGGCC 720
QY      241 GlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuAlaArgProGlySerSerLeuArg 260
DB      721 CAGCTCGCGCTGTGCGCAGAGGTGTGGCTGTGGCCCTGCGGCGCAGGTCCTCCGCG 780
QY      261 IleArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTrrPheGlyLeu 280
DB      781 ATCCGACACCTCCCGCGCGCATCTCAAGAGCGGCCCTTCTCTACATTCGAGATC 840
QY      281 PheGlnValHis 284
DB      841 TTCAGAGTTCAC 852
XX
XX      RESULT 2
XX      ABRK34881
XX      ID      ABRK34881 standard; cDNA; 1364 BP.
XX      AC      ABRK34881;
XX      DT      08-MAY-2002 (first entry)
XX      DE      Human cDNA encoding secreted protein #19.
XX      KW      Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
XX      viral infection; bacterial infection; fungal infection; diabetes; asthma;
XX      autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
XX      autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
XX      Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
XX      coagulation disorder; inflammatory disorder; Crohn's disease; incision;
XX      tissue regeneration; wound healing; burn; haematopoiesis;
XX      myeloid cell deficiency; lymphoid cell deficiency.
XX      OS      Homo sapiens.
XX      PN      WO200177288-A2.
XX      PD      18-OCT-2001.
XX      PF      29-MAR-2001; 2001WO-US10224.
XX      PR      06-APR-2000; 2000US-195582P.
```

XX (GENY ) GENETICS INST INC.  
 PA  
 XX Wong GG, Clark HF, Rechele K, Agostino MJ, Howes SH, Resnick RJ;  
 PI Gulukota K, Graham JR;  
 XX WPI: 2002-179321/23.  
 DR  
 XX  
 PT Five hundred and ninety two polynucleotides derived from a variety of  
 PT human tissue sources which encode secreted proteins, useful for  
 PT treating immune deficiencies and disorders such as autoimmune disorders  
 PT  
 PS  
 PS  
 PS  
 CC Claim 1; Page 82; 372pp; English.  
 CC  
 CC The invention relates to 592 polynucleotides which have been derived from  
 CC a variety of human tissue sources and which encode novel secreted  
 CC proteins. The polynucleotides can be used as probes for the  
 CC identification and isolation of full length cDNA and genomic DNA. The  
 CC polynucleotides and proteins can also be used as nutritional supplements.  
 CC The proteins are useful in the treatment of various immune deficiencies  
 CC and disorders such as viral infections, bacterial infections, fungal  
 CC infections, autoimmune disorders (e.g. Rheumatoid arthritis, multiple  
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
 CC and conditions (e.g. asthma). They are also useful for treating  
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
 CC disease), liver fibrosis, coagulation disorders (e.g. hemophilia),  
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
 CC useful for tissue regeneration, for wound healing and in the treatment of  
 CC burns, incisions and ulcers. The proteins are also useful for regulating  
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
 CC Sequences ABR34863-ABR35454 represent polynucleotides of the invention.  
 CC  
 XX  
 SO Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 5,08e-93 Length: 1364  
 Score: 1406.00 Matches: 276  
 Percent Similarity: 100.00% Conservative: 0  
 Best local Similarity: 100.00% Mismatches: 0  
 Query Match: 97.37% Indels: 0  
 DB: Gaps: 0

US-09-245-198A-4 (1-284) x ABR34861 (1-1364)

OY 9 SerAlaAargyleuProleuProargSerleuGlySerArgAspGlyAlaValArg 28  
 DB 5 TCCGCGCCGCGGCTCCCTCCCGATCCCTCGGGTCCCGGAGGGGGCGGTGAGG 64  
 OY 29 GlnAlaGlnProProAlaProMetaAlaAlaAargSerGlnAargArgGlyArgArg 48  
 DB 65 CAGGACAGACCCCGCCCGCCATGGCCCGCGTGGAGACCGAGAGCGAGGGGCGCCGG 124  
 OY 49 GlyGluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAla 68  
 DB 125 GGGGAGCGGGGACCGCCTGCTGCTGCCGCTCGCGCTGGCGCTGGCGCTGGCC 184  
 OY 69 CysLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGln 88  
 DB 185 TGCCTCGGCGCTCTGCTGGCGGTGCTGAGTTGGGAGCCGGGCACTCCCTGCCCCAG 244  
 OY 89 GluProAlaGlnGluGluLeuValAlaGluGluAspGlnAspProSerGlyLeuLeuAsnPro 108  
 DB 245 GAGCTGCGCCAGAGAGACTGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 304  
 OY 109 GlnThrGluGluSerGlnAspProAlaProPheLeuAsnAargLeuValAargProAargArg 128  
 DB 305 CAGACAGAAAGAGAGAGATCTCGCTTCTCGAACCAGCATGTTCTGGCGCTGCAGAG 364  
 OY 129 SerAlaProGlyGlyArgGlyThrArgAlaAargArgAlaAlaAlaAlaAlaAlaAlaAla 148  
 DB 365 AGTGCACCTAAAGCGCGGAAAAACAGAGGCTCGAAGAGGATCGAGCCCATTTATGAAGTT 424

OY 149 HisProAargProGlyGlnAspGlyAlaGlnAlaValAlaAspGlyThrValSerGlyTrp 168  
 DB 425 CATCCAGACCTCGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484  
 OY 169 GluGluAlaAargGlyLeuAsnSerSerSerProLeuAargTrpAsnAargGlnIleGlyGluPhe 188  
 DB 485 GAGAGAGCCAGAAATCAACAGCTCCAGCCCTGCGCTCAACACCGCCAGATCGGAGATT 544  
 OY 189 IleValThrAargAlaGlyLeuThrTyrrLeuTyrrGlyGlnValHisPheAspGlnGlyLys 208  
 DB 545 ATAGTCACCCGCGGCTGCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604  
 OY 209 AlaValTyrrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuAlaLeuAlaLeuAla 228  
 DB 605 GCTGCTACCTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664  
 OY 229 GluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuAargLeuGlyGlnValSer 248  
 DB 665 GAATCTCAGCCACATGCGGAGATTCCCTCGGGCCCGAGCTCCGCTGCGCAGGTGCT 724  
 OY 249 GlyLeuLeuAlaLeuAargProGlySerSerLeuAargIleAargThrLeuProThrAlaHis 268  
 DB 725 GGGCTGTGGCCCTGCGGCGAGGCTCTCCCTGCGATCCGACCCCTCCCTGGGCCAT 784  
 OY 269 LeuLysAlaAlaProPheLeuThrTyrrPheGlyLeuPheGlnValHis 284  
 DB 785 CTCAGAGCTGCCCTTCTCTCACTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCTTCT 832  
 RESULT 3  
 AAX56000  
 ID AAX56000 standard; DNA; 1421 BP.  
 XX  
 AC AAX56000;  
 XX  
 DT 15-JUL-1999 (first entry)  
 XX  
 DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.  
 XX  
 KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;  
 KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
 KW cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 92..841  
 FT /\*tag= a  
 FT /product= "Apo-3 ligand"  
 XX  
 XX MO9919490-A1.  
 XX  
 PD 22-APR-1999.  
 XX  
 XX 09-OCT-1998; 98WO-US21407.  
 XX  
 XX 17-DEC-1997; 97US-0069862.  
 PR 10-OCT-1997; 97US-0062037.  
 XX  
 XX (GENY ) GENENTECH INC.  
 PA  
 PI Ashkenazi AJ, Marsters SA, Pittl R;  
 XX WPI: 1999-287982/24.  
 DR P-PSDB; AAY09369.  
 XX  
 XX New human Apo-3 ligand (a tumour necrosis factor) homologue  
 XX  
 PS Claim 18; Fig 1; 74pp; English.  
 CC  
 CC The present sequence encodes a human tumour necrosis factor (TNF) and  
 CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
 CC cytosolic activity. Apo-3 ligand can be used to induce apoptosis in  
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and



Score: 1366.00 Matches: 268  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 94.60% Indels: 0  
 DB: 21 Gaps: 0

US-09-245-198a-4 (1-284) x AAA49717 (1-1353)

```

Oy 17 ArgSerLeuGlySerArgAspGlyGlyAlaValAlaArgInAlaGlnProProAlaProMet 36
Db 1 CGATCCCTCGGGTCCCGGATGGGGGGGGGTGAGGACAGACACCCCGCCCGCCATG 60
Oy 37 AlaAlaArgSerGlnArgArgArgGlyArgArgGlyGlnProGlyThrAlaLeuLeu 56
Db 61 GCGCGCCGTCGGAGGACAGAGGGGAGGGGCGCCCGGGGGAGACCCGCGCTGCTG 120
Oy 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaVal 76
Db 121 GTCCCGCTCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 180
Oy 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGlnProAlaGlnGlnGlnGlnVal 96
Db 181 GTCACTTTGGGAGCGGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Oy 97 AlaGlnGlnAspGlnAspProSerGlnLeuAsnProGlnThrGlnGlnSerGlnAspPro 116
Db 241 CGAGAGGAGGACAGACCGCCGTCGGAACATGCCACAGAGAGAGAGAGAGATCCT 300
Oy 117 AlaProPheLeuAsnArgLeuValAlaArgProArgArgSerAlaProGlyGlyArgIleThr 136
Db 301 GCGCTTCTCTGTAACGCACTAGTCTGCGCTGCGCAGAGAGTGCACCTAAAGGCGGAGAAACA 360
Oy 137 ArgAlaArgArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 156
Db 361 CGGGCTCGAAGAGCATGCGACCGCATTTATGAGTTATCATCCAGACCTGAGAGAGCGGA 420
Oy 157 AlaGlnAlaGlyValAlaAspGlyThrValSerGlyTyrGlnGlnAlaArgIleAsnSerSer 176
Db 421 GCGGAGGAGGAGTGTGGAGGAGGAGTGTGGAGTGTGGAGGAGGAGGAGGAGTGTGC 480
Oy 177 SerProLeuArgTyrAsnArgGlnIleGlyGlnPheIleValThrArgAlaGlyLeuTyr 196
Db 481 AGCCCTGTCGCTACACCGCAGATCGGGAGTTTATAGTCAACCGGCTGCGCTCTAC 540
Oy 197 TyrLeuTyrCysGlnValAlaHisPheAspGlnGlyLysAlaValTyrLeuLeuAspLeu 216
Db 541 TACCTGTACTGTGACAGTGCACCTTTATGATGAGGAGGAGGAGGAGGAGGAGGAGTGTG 600
Oy 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGlnGlnGlnPheSerAlaThrAlaAlaSer 236
Db 601 CTGGTGAGATGGTGTGCTGCGCCCTGCGCTGCGGAGGAGGAGGAGGAGGAGGAGTGTG 660
Oy 237 SerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGly 256
Db 661 TCCCTCGGAGCCCGACCTCGCTGCGCAGGTGTGCTGCGCTGTGCGCTGTGCGCGGAGG 720
Oy 257 SerSerLeuArgIleArgThrLeuProPheAlaHisLeuValAlaAlaProPheLeuThr 276
Db 721 TCTCTCCCTCGGATCGGACACCTCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGTGTG 780
Oy 277 TyrPheGlyLeuPheGlnValHis 284
Db 781 TACTTCGAGCTCTTCACAGTTTAC 804

```

RESULT 5

ABK40255

ID ABK40255 standard: cDNA; 1353 BP.

AC ABK40255;

DE 15-JUL-2002 (first entry)

CDNA encoding human PRO207 polypeptide.

```

XX Human: PRO: benign tumour; malignancy; lymphoid malignancy;
KW Leukaemia; neuronal disorder; stromal disorder; blastocoele disorder;
KW Inflammatory disorder; immune disorder; angiogenic disorder;
KW gene therapy; cytosolic; neuroprotective; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200153486-A1.
PD 26-JUL-2001.
XX
XX 11-FEB-2000; 2000WO-US03565.
XX
XX 08-MAR-1999; 99WO-US05028.
XX 11-MAR-1999; 99US-123972P.
XX 11-MAY-1999; 99US-133459P.
XX 02-JUN-1999; 99WO-US12252.
XX 22-JUN-1999; 99US-140650P.
XX 22-JUN-1999; 99US-140653P.
XX 20-JUL-1999; 99US-144758P.
XX 26-JUL-1999; 99US-145698P.
XX 28-JUL-1999; 99US-146222P.
XX 17-AUG-1999; 99US-149395P.
XX 31-AUG-1999; 99US-151689P.
XX 01-SEP-1999; 99WO-US20111.
XX 15-SEP-1999; 99WO-US21090.
XX 30-NOV-1999; 99WO-US28313.
XX 01-DEC-1999; 99WO-US28301.
XX 01-DEC-1999; 99WO-US28634.
XX 05-JAN-2000; 2000WO-US00219.
XX
PA (GENE) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Masters SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
DR WPI: 2002-205567/26.
DR P-PSDB; AAU86129.
XX
XX Thirty five nucleic acids encoding PRO polypeptides, useful for
XX treating benign or malignant tumours, leukaemias and lymphoid
XX malignancies, inflammatory, angiogenic and immunologic disorders -
XX
XX Claim 50; Fig 3; 302pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
XX polypeptides and the polynucleotide sequences encoding them. The
XX CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
XX useful for treating benign or malignant tumours (e.g. renal, kidney,
XX bladder, breast, etc), leukaemias and lymphoid malignancies, other
XX disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
XX CC macrophagal, stromal and blastocoele disorders, inflammatory, immune
XX CC and angiogenic disorders. The polynucleotide sequences are also
XX CC useful in gene therapy. ABK40254-ABK40288 encode for the human PRO
XX CC polypeptides of the invention.
XX
SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other:

```

Alignment Scores:

Pred. No.: 3.96e-90 Length: 1353  
 Score: 1366.00 Matches: 268  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 94.60% Indels: 0  
 DB: 24 Gaps: 0

US-09-245-198a-4 (1-284) x ABK40255 (1-1353)

```

Oy 17 ArgSerLeuGlySerArgAspGlyGlyAlaValAlaArgInAlaGlnProProAlaProMet 36
Db 1 CGATCCCTCGGGTCCCGGATGGGGGGGTGAGGACAGACACCCCGCCCGCCATG 60

```

QY 37 AAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeu 56  
 |||  
 Db 61 GCCGCCGCTCGAGCAGAGCGAGGCGCGCGGAGGCGCGGCGCTGCTG 120  
 QY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 76  
 |||  
 Db 121 GTCCCGCTCGCGGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 180  
 QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96  
 |||  
 Db 181 GTGAGTTTGGGGAGCCCGGATCGCTGCTGCGCAGAGAGCTTCCCAAGGAGAGCTG 240  
 QY 97 AlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAspPro 116  
 |||  
 Db 241 GCAGAGAGAGACAGACCGCTGGAACCTGAATCCCGACAGAGAAAGACAGATCCT 300  
 QY 117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 136  
 |||  
 Db 301 GCCCTTTCTGAAACCGACTAGTTGCGCTCGCAGAAAGTCCACCTAAAGCGCGAAACA 360  
 QY 137 ArgAlaArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 156  
 |||  
 Db 361 CGGGCTCGAAGAGCGATCGCAGCCATTATGAGATTATCCAGAGACCTGAGACAGGGA 420  
 QY 157 AlaGlnAlaGlyValAlaAspGlyThrValSerGlyTyrPoliGluAlaArgIleAsnSerSer 176  
 |||  
 Db 421 GCCAGAGCAGAGTGGAGCGGAGACAGTGGCTGGGAGAGACCAAGATCAACAGCTCC 480  
 QY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196  
 |||  
 Db 481 AGCCCTCTGGCTGACAAACCCAGATCGGGAGATTATAGTTCACCCGGGCTGGCTCTAC 540  
 QY 197 TyrLeuTyrCysGlnValHisPheAspGlyGlyValAlaValTyrLeuLysLeuAspLeu 216  
 |||  
 Db 541 TACCTGACTGATGAGGCTGCTGATGAGGAGGAGGCTCTCACTGAAAGCTGGACTTG 600  
 QY 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaIleSer 236  
 |||  
 Db 601 CTGGTGAGATGCTGCTGGGCGCTGCGCTGCTGAGGAATTTTCAGCACTCGCGCAGT 660  
 QY 237 SerLeuGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256  
 |||  
 Db 661 TCCCTGGGCGCCAGCTCGCGCTCTGCAAGTGTGGCTGTGGCGCTGCGCGCAGGG 720  
 QY 257 SerSerLeuArgIleArgThrLeuProTyrPalaHisLeuLysAlaIleAlaProPheLeuThr 276  
 |||  
 Db 721 TCCCTCCCTGGGATCGGCACTCTCCCTGGGCGCCATCTCAAGCTGCGCTTCTTCACC 780  
 QY 277 TyrPheGlyLeuPheGlnValHis 284  
 |||  
 Db 781 TACTTCGACTCTTCCAGGTTTCAC 804  
 RESULT 6  
 AAV47613  
 ID AAV47613 standard; cDNA: 1236 BP.  
 AC AAV47613;  
 XX  
 XX 27-OCT-1998 (first entry)  
 DT  
 XX TNF related endothelium proliferative agent gene.  
 DE  
 XX ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
 KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..750  
 CDS /tag= a  
 FT /product= "TREPA"  
 FT  
 XX

PN W09835061-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 12-FEB-1998; 98WO-US02859.  
 XX  
 PR 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI; 1998-447255/38.  
 DR P-PDB; AAM29745.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 treatment of autoimmune disease, tumours and inflammation  
 PS Claim 11; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 XX  
 SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 7, 37e-83 Length: 1236  
 Score: 1265.00 Matches: 248  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.60% Mismatches: 0  
 Query Match: 87.60% Indels: 0  
 DB: 19 Gaps: 0  
 US-09-245-198a-4 (1-284) x AAV47613 (1-1236)  
 QY 36 MetaAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeu 55  
 |||  
 Db 1 ATGGCGCGCGCGTGGAGCCAGAGCGAGGCGCGCGGAGGAGCGCGGCGCTG 60  
 QY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 75  
 |||  
 Db 61 CTGGTCCCGCTCGCTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 120  
 QY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 95  
 |||  
 Db 121 GTGCTAGTTTGGGGAGCCCGGATCGCTGCTGCGCAGAGAGCTTCCCAAGGAGAGCTG 180  
 QY 96 ValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAsp 115  
 |||  
 Db 181 GTGGCAGAGAGAGACAGACCGCTGGAACCTGAATCCCGACAGAGAAAGACAGAT 240  
 QY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 135  
 |||  
 Db 241 CCGCGCGCTTCTGAAACCGACTAGTTGCGCTCGAAGAAAGTCCACCTAAAGCGCGAAA 300  
 QY 136 ThrArgAlaArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 155  
 |||  
 Db 301 ACAGGGGCTCGAAGAGCGATCGCAGCCATTATGAGATTATCCAGAGACTGGAGACGAC 360



QY 156 GYAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPGluGluAlaArgIleAsnSer 175  
 Db 361 GGAGCGCAGGAGGTGTGACGGGACACTGATGCTGGAGGAGCAAGATCAACAGC 420  
 QY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195  
 Db 421 TCCAGCCCTTCGCGTACACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTC 480  
 QY 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAsp 215  
 Db 481 TACTACCTGTACTGTACAGTGCACCTTTATGAGGGGAGGAGCTGTACCTGAGGCTGGAC 540  
 QY 216 LeuLeuValAspGlyValIleAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235  
 Db 541 TTGCTGGTGGATGGTGTCTGGCCCTGGCTGGCTGGAGGAATCTCAAGCAGTGGCGG 600  
 QY 236 SerSerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgPro 255  
 Db 601 AGTTCCCTCGGGCCCGACCTCGGCTGCGCAGGAGTGTGGCTGTGGCCCTGGCGGCA 660  
 QY 256 GlySerSerLeuArgIleArgThrLeuProTrpAlaHisIleLysAlaAlaProPheLeu 275  
 Db 661 GGGTCTCTCCGCGGATCCGACACCTCCCTGGGCCATCTCAAGGCTGCCCTTCTCTC 720  
 QY 276 ThrTyrPheGlyLeuPheGlnValHis 284  
 Db 721 ACCTACTCTCGGACCTCTCCAGGTTCCAC 747

RESULT 7  
 AAD04350  
 ID AAD04350 standard; cDNA; 1236 BP.  
 AC AAD04350;  
 XX 04-JUN-2001 (first entry)  
 DE Human TREPA (TNF related endothelium proliferative agent) cDNA.  
 XX Human; tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vulnetary; ss.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 1..750  
 FT /tag= a  
 FT /product= "Human TREPA (TNF related endothelium  
 FT proliferative agent)"  
 FT  
 XX US6207642-B1.  
 PN 27-MAR-2001.  
 PD 26-JUN-1998; 98US-0105343.  
 XX 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX (ABBO ) ABBOTT LAB.  
 PA  
 XX WAlley SR;  
 PI WPI: 2001-280760/29.  
 DR P-PSDB: AAE00891.  
 XX Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 XX Example 2; Column 73-74; 53pp; English.  
 PS

CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent),  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present sequence is a cDNA clone ID #690050 encoding human TREPA.  
 CC  
 XX  
 SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 7.37e-83 Length: 1236  
 Score: 1265.00 Matches: 248  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.60% Mismatches: 0  
 Query Match: 87.60% Indels: 0  
 DB: Gaps: 0  
 US-09-245-198A-4 (1-284) x AAD04350 (1-1236)  
 QY 36 MetaAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeu 55  
 Db 1 ATGGCCGCCCGTCGGAGCAGCAAGCGAGGGGCGCCGGGGGAGCCGGGCAACCCCTG 60  
 QY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75  
 Db 61 CTGGTCCCGCTCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 120  
 QY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeu 95  
 Db 121 GTGGTCAGTTTGGGAGCGCGGCGATCGCTGTCCGCCAGAGCTGCCAGAGGAGCTG 180  
 QY 96 ValAlaGluGluAspGluAspProSerGluLeuAsnProGlnInThrGluGluSerLys 115  
 Db 181 GTGGCAGAGGAGGACGAGGACCGGTCGGAACGTAATCCCAACAGCAAGAAAGCCAGAT 240  
 QY 116 ProAlaProPheLeuAsnArgLeuValArgProAlaArgSerAlaProLysGlyArgLys 135  
 Db 241 CCTGGCCTTTCCCTAACAGCACTAGTTGGGCTCGAAGAGTGCACCTTAAGCGCGGAAA 300  
 QY 136 ThrArgAlaArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAsp 155  
 Db 301 ACACGGGCTCGAAGAGCATGCGACGCCCATATGATGATTCACGAGCTGGACAGGAC 360  
 QY 156 GYAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPGluGluAlaArgIleAsnSer 175  
 Db 361 GGAGCGCAGGAGGTGTGACGGGACAGTGAAGTGTGGAGGAGCAAGATCAACAGC 420  
 QY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195  
 Db 421 TCCAGCCCTTCGCGTACACCGCCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTC 480  
 QY 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAsp 215  
 Db 481 TACTACCTGTACTGTACAGTGCACCTTTATGAGGGGAGGAGCTGTACCTGAGCTGGAC 540  
 QY 216 LeuLeuValAspGlyValIleAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235  
 Db 541 TTGCTGGTGGATGGTGTCTGGCCCTGGCTGGCTGGAGGAATCTCAAGCAGTGGCGG 600  
 QY 236 SerSerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgPro 255  
 Db 601 AGTTCCCTCGGGCCCGACCTCGGCTGCGCAGGAGTGTGGCTGTGGCCCTGGCGGCA 660  
 QY 256 GlySerSerLeuArgIleArgThrLeuProTrpAlaHisIleLysAlaAlaProPheLeu 275  
 Db 661 GGGTCTCTCCGCGGATCCGACACCTCCCTGGGCCATCTCAAGGCTGCCCTTCTCTC 720  
 QY 276 ThrTyrPheGlyLeuPheGlnValHis 284  
 Db 721 ACCTACTCTCGGACCTCTCCAGGTTCCAC 747

RESULT 8  
 AAS03964  
 ID AAS03964 standard; DNA; 898 BP.  
 XX  
 AC AAS03964;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE Expression vector pDC409-L2-TWEAK fusion protein-encoding DNA.  
 XX  
 KM TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KM rubecsis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 KM corneal graft neovascularisation; psoriasis; metastatic condition;  
 KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KM preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KM peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 KM fusion protein.  
 XX  
 OS Homo sapiens.  
 XX Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT 52..873 /tag= a  
 FT CDS /product= "Fusion protein comprising a growth hormone  
 FT leader, a leucine zipper multimerisation  
 FT domain, and human TWEAK extracellular  
 FT domain"  
 XX  
 PN WO200145730-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 19-DEC-2000; 2000WO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMMV ) IMMUNEX CORP.  
 XX  
 PI Wiley SR;  
 DR WPI: 2001-417975/44.  
 DR P-PSDB; AAU034499.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 XX  
 PS Example 1; Page 39-40; 46pp; English.  
 XX  
 CC The sequence represents a DNA from the expression vector  
 CC pDC409-L2-TWEAK, which encodes a fusion protein comprising a growth  
 CC hormone leader, a leucine zipper multimerisation domain, and the  
 CC extracellular domain of human TWEAK. The fusion protein was used in  
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 CC from a COS cell human cDNA library. The TWEAK protein is a  
 CC member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 CC used in the treatment and diagnosis of human disease. The disorders  
 CC mediated by angiogenesis include ocular disorders characterised by ocular  
 CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,  
 CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
 CC rubecsis, uveitis, macular degeneration and corneal graft  
 CC neovascularisation, and inflammatory diseases such as arthritis,  
 CC rheumatism and psoriasis. Other treatable diseases include malignant and  
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and

CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
 CC scleroderma, vascular adhesions, atherosclerotic plaque  
 CC neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 XX  
 SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,32e-68 Length: 898  
 Score: 1066.00 Matches: 207  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 73.82% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-245-198a-4 (1-284) x AAS03964 (1-898)  
 QY 78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuValAla 97  
 DB 250 AGTTTGGGAGCCGGGCAATCGCTGCCGCCAGAGCCTGCCAGAGAGAGTGGTGCA 309  
 QY 98 GluGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAspProAla 117  
 DB 310 GAGGAGACCCAGACCCGCTCGAATGAAATCCCAAGCAAGAAAGCCAGATCTGCG 369  
 QY 118 ProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThrArg 137  
 DB 370 CTTTCTCGAAGCGACTAGTGGCCCTCGAGAAAGACACTAAAGCCGGAAACACGG 429  
 QY 138 AlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAla 157  
 DB 430 GCTCGAAGACGATCGCAGCCGCAATTATGAAGTTCAATCCAGACCTGGACAGCGAGCG 489  
 QY 158 GlnAlaGlyValAspGlyThrValSerGlyTyrPGLuGlnAlaArgIleAsnSerSer 177  
 DB 490 CAGGCAAGTGTGGACGGGACAGTGAAGTGGCTGGAGAGAACCCAGATCAACAGCTCCAC 549  
 QY 178 ProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyr 197  
 DB 550 CCTCGGGCTACAAACCGCCGAGTCGGGAGTTATATGATCAACCCGGGCTGGCTACTAC 609  
 QY 198 LeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAspLeu 217  
 DB 610 CTGTACTGTGAGGTGCACATTGATGAGGGAAGGCTGTCTACTGGAAGCTGACTGTG 669  
 QY 218 ValAspGlyValLeuAlaLeuAlaArgCysLeuGluGlnPheSerAlaThrAlaAlaSer 237  
 DB 670 GTGGATGGTGTGCTGGCCCTGGCTGCTGGAGAAATTCACACCACTGGGGCAATTCC 729  
 QY 238 LeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySer 257  
 DB 730 CTCGGGCCCCAGCTCCGCTCTGCCAGGTGTGGGCTGTGGCCCTGGCGCCAGGGTCC 789  
 QY 258 SerLeuArgIleArgThrLeuProTrrAlaHisLeuLysAlaIleAlaProPheLeuThrTyr 277  
 DB 790 TCCCTGGGATCCGACACCTCCCTCCCGCCATCTCAAGGCTGCCCTTCCTCACCTTAC 849  
 QY 278 PheGlyLeuPheGlnValHis 284  
 DB 850 TTGGGACTCTCCAGGTTTCAC 870  
 RESULT 9  
 AAX23424  
 ID AAX23424 standard; DNA; 1030 BP.  
 XX  
 AC AAX23424;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Human TNRL3 DNA.  
 XX  
 KM Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;  
 KM developmental abnormality; gestational abnormality; prostate cancer;

KW	AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis: treatment; therapy; disease;
KW	cycloplasmic domain; immunogen; antibody preparation; Breast carcinoma;
KW	apoptosis; human; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	location/Qualifiers
FT	CDS
FT	1..627
FT	/*tag= a
FT	/product= "TNRL3"
XX	
PN	W09911791-A2.
XX	
PD	11-MAR-1999.
XX	
PE	04-SEP-1998; 98WO-US18393.
XX	
PR	05-SEP-1997; 97US-0924634.
XX	
PA	(UNIW ) UNIV WASHINGTON.
XX	
PI	Chaudhary PM;
XX	
DR	WPI; 1999-205191/17.
DR	P-PSDB; AAW93590.
XX	
PT	New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT	useful for diagnosis and treatment of prostate cancer and
PT	developmental or gestational abnormalities
XX	
SS	Example VII; Fig 13A; 156pp; English.

CC This invention describes isolated Tumor Necrosis factor (TNF) family  
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
CC fragments, and isolated TNF related ligands 1 and 3 (TNR1L and TNR13) or  
CC their active fragments. APO4 is useful for diagnosing prostate cancer  
CC by determining levels of APO4 in an individual. Prostate cancer can also  
CC be treated using APO4 selective binding agents linked to a therapeutic  
CC moiety. APO4 polypeptides are also useful for identifying selective  
CC binding agents, useful in diagnosis/treatment of disease by binding of  
CC agents to the polypeptide/active fragment which is extracellular, or  
CC expressed on the cell surface. The binding is preferably performed in  
CC vivo. APO4 polypeptides/active fragments are also useful for screening  
CC for agonists and antagonists by binding and observing the change in APO4  
CC activity. Effective pharmacological agents useful in diagnosis or  
CC treatment of disease are also identified using APO4 polypeptides/active  
CC fragments and APO4 signal transducer molecules that specifically interact  
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
CC activity. The method is performed in vivo or in vitro. APO polypeptides  
CC are all useful as immunogens for preparing antibodies. APO4 is also  
CC useful for diagnosis/treatment of developmental or gestational  
CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
CC MCF-7, and induced apoptosis.

XX

SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Alignment Scores:

Prid. No.:	3 01e-68	Length:	1030
Score:	1062.00	Matches:	207
Percent Similarity:	99.52%	Conservative:	0
Best Local Similarity:	99.52%	Mismatches:	1
Query Match:	73.55%	Indels:	0
AB:	20	Gaps:	0

US-09-245-198A-4 (1-284) x AAX234z24 (1-1030)

OY ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96  
Dbb 1 GTCAAGTTGGGGAACCGGCATCGTCGCCACAGAGACTTGCCACAGAGAGCTGTG 60

Oy 97 AlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro 116  
Db 61 GCACAGGAGGACACGAGCAGCCGTCGGAATGTAATCCCACACAGAGAAGAACGAGCATCT 120

QY	117	AlaPhePheLeuAsnArgLeuValArgProIargrSerAsnAlaProlySglAryGlyThr	136
Db	121	GGCGCTTCCTGAAACCGACTAGTGGCTGGCAGAAAGTGCATTAAGGCCGGAACCA	190
QY	137	ArgAlaArgAlaAlaIleAlaAlaHisIstGrluValHisProArgProGlyGlnAspGly	156
Db	181	CGGGCTCGAAGAGCATGCGAGCCCATTTATGAATTCTACACAGCTGGACAGACGGA	240
QY	157	AlaGlnAlaGlyValAspGlyThrValSerGlyTyrPgluGlnAlaArgIleasnSer	176
Db	241	GGCGAGCAGAGGTGTGGAGCGAGCAGTGTGGTGGGGAGAACCAATCAACAGCTCC	300
QY	177	SerProIleuArgTyrAsnArgGlnIleGlyIupheIleValIThrArgAlaGlyLeuTyr	196
Db	301	AGCCCTCTGCGGTACTACCGCAGATCGGGAGTTTATAGTCACCCGGGCTGGCTTAC	360
QY	197	TyrLeuTyrCysGlnValHisPheAspGluIuLySAlaValTyrLeuLySLeuAspLeu	216
Db	361	TACCTGTACTGTACAGGTGCACCTTTGATGAGGGAGGAAGGTGTCATCCGAAAGTGGACTTG	420
QY	217	LeuValAspGlyValLeuAlaLeuAlaArgCysLeuGluGluIupheSerAlaThrAlaIaSer	236
Db	421	CTGGTGGATGTGTGCTGTGCCCTGGCGCTCGAGGAAATTCATCAGCCACTGGCGGCAGT	480
QY	237	SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGly	256
Db	481	TCCCTCGGGCCCCAGCTCGGCTGTGCCAGGTGTCGTGGCTGTGGCCCTGGCGCAGGG	540
QY	257	SerSerLeuArgIleArgThrLeuProTrrAlaHisLeuLySAlaAlaIaProPheLeuThr	276
Db	541	TCCCTCGGCGGATGCCGAGCCCTCCCTGGGGCCCATGTCAAGGCTGGCCCTTCTCACC	600
QY	277	TyrPheGlyLeuPheGlnValHis	284
Db	601	TACTTCGGACTCTTCCAGTTCAC	624
RESULT 10			
AAV18599			
ID	AAV18599 standard; cDNA; 1168 BP.		
AC	AAV18599;		
XX			
DT	21-JUL-1998 (first entry)		
XX			
DE	Mus musculus tumour necrosis factor related ligand (TRELL) gene.		
XX			
KW	TRELL; tumour necrosis factor related ligand; tnf; treatment;		
KM	cancer; autoimmune disease; immune system; stimulation; suppression;		
KW	graft rejection; ds.		
XX			
OS	Mus musculus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	2..679	
FT		/*tag= a	
FT		/note= "tumour necrosis factor related ligand"	
PN	W09805783-A1.		
XX			
PD	12-FEB-1998.		
XX			
PF	07-AUG-1997; 97MO-US13945.		
XX			
PR	18-MAR-1997; 97US-0040820.		
PR	07-AUG-1996; 96US-0023541.		
PR	18-OCT-1996; 96US-0028515.		
XX			
PA	(BIOJ ) BIOGEN INC.		
XX			
PA	(UYGE-) UNIV GENEVA FACULTY MEDICINE.		
XX			
PI	Browning JL, Chicheportiche Y;		

[illegible]

DR WPI: 1998-145619/13.  
 DR P-PSDB; AAW47524.  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 XX  
 PS Claim 2; Pages 45-46; 69pp; English.  
 XX  
 CC The sequence is that encoding mouse tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. The DNA sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of probe probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 CC  
 XX  
 XX Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 3.8e-65 Length: 1168  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 70.64% Indels: 0  
 Gaps: 0  
 DB:  
 US-09-245-198a-4 (1-284) x AAW18599 (1-1168)  
 QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValaSerLeuGly 80  
 DB 5 CTGAGCCTCGGGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGG 64  
 QY 81 SerArgLaserLeuSerArgLaserArgLaserArgLaserArgLaserArgLaser 100  
 DB 65 AGCTGGGCAAGCTGCTGCTGGCCAGAGAGCTTCAGAGAGAGAGAGAGAGAGAGAG 124  
 QY 101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeu 120  
 DB 125 CGGAGGCCCTCGAATGAAATCCCAAGACAGAGAAAGCCAGATGCTGTTCTTG 184  
 QY 121 AsnArgLeuValArgProArgSerArgProArgSerArgSerArgSerArgSerArg 140  
 DB 185 GAAACAATACCTCGCGCTCGAAGAAGTCTCTTAAGCCGGGAAGGGCGCTCGCGCA 244  
 QY 141 AlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
 DB 245 GCTATTTCAGACCATTAATGAGGTTCATCTCGGCCAGACAGATGAGACAAAGAGT 304  
 QY 161 ValAspGlyThrValSerGlyTTPGluGluAlaAlaArgIleAsnSerSerProLeuArg 180  
 DB 305 GTGAGTGCGGACAGTGAAGTCTGCTGGAGAGCAACCAAAATCAACAGCTCCACCTCTCGC 364  
 QY 181 TyrAsnArgLInIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
 DB 365 TACGACCGCCAGATGGGGAATTTACAGTCATCAGGGCTGGCTGACTACTACTGT 424  
 QY 201 GlnValHisPheAspGlyGlyValAlaValTyrLeuLeuAspLeuLeuValAspGly 220  
 DB 425 CAGGTGACATTGATGAGGAAAGAGCGTCTACTCTGAAGCTGAGCTTGGTGAACGGT 484  
 QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerSerLeuGlyPro 240

DB 485 GTCTGGCCCTCGCGCTCGGAAATTCAGCCACACAGACAGCTCTCCGCGCCC 544  
 QY 241 GlnLeuArgLeuGlyGlnValSerGlyLeuAlaLeuArgProGlySerSerLeuArg 260  
 DB 545 CAGCTCCGTTTGGCCAGGAGTCTGCGCTGGCGGCGGAGGTTCCCTTCGG 604  
 QY 261 IleArgThrLeuProThrAlaHisLeuAlaAlaProPheLeuThrTyrPheGlyLeu 280  
 DB 605 ATCCGACCCCTCCCGTGGCTCATCTTAAGCTGCCCCCTTCCAACTTCTTGACCTC 664  
 QY 281 PheGlnValHis 284  
 DB 665 TTTCAGACTCAC 676  
 RESULT 11  
 ID AAX23425 standard; DNA: 701 BP.  
 AC AAX23425;  
 DT 18-JUN-1999 (first entry)  
 XX Mouse TNRL3 DNA.  
 XX  
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse; ss.  
 OS Mus sp.  
 XX  
 XX Key Location/Qualifiers  
 FH 1.636  
 FT CDS /\*tag= a "TNRL3"  
 FT /product= "TNRL3"  
 FT  
 XX WC0911791-A2.  
 XX  
 XX PD 11-MAR-1999.  
 XX  
 XX PF 04-SEP-1998; 98WO-US18393.  
 XX  
 XX PR 05-SEP-1997; 97US-0924634.  
 XX  
 XX PA (UNITW ) UNIV WASHINGTON.  
 XX  
 XX PI Chaudhary PW;  
 XX  
 XX DR WPI: 1999-205191/17.  
 DR P-PSDB; AAW93591.  
 XX  
 XX PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 XX PS Example VII; Fig 13B; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active

CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.

SO Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

#### Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:
Score: 968.00	701	188	9
Percent Similarity: 93.36%			
Best Local Similarity: 89.10%		Mismatches: 14	
Query Match: 67.04%		Indels: 0	
	Gaps: 0		

US-09-245-198A-4 (1-284) x AAV23425 (1-701)

```

OY 74 LeuAlaValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CTGCTCGTGTGAGCTGTGGAGCTGAGCAACGCTGTCTGCCAGACCTTCTCAGAG 60

OY 94 GluLeuValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSer
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GAGCTGACAGCAGAGAGACCGCGGAGCCCTGACACTGATATCCCAACAGAGAAAGC 120

OY 114 GlnAspProAlaProPheLeuAsnArgLeuValArgProArgSerAlaProGlySgly
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 CAGGATGTGTACTTCTTGGAACTAGTCCGCGCTCGAAGAAAGTCTCTAAAGC 180

OY 134 ArgLysThrArgAlaArgArgAlaAlaAlaHisTyrGluValHisProArgProGly
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CGGAAGCGCGCGCTCGCGGAGCTATTACGCCCTTGTGAGTTCATCTCGGACAGA 240

OY 154 GlnAspGlyAlaGlnAlaGlyValAlaAspGlyThrValSerGlyTTPGlnGluAlaArgIle
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 CAGGTGAGCACAAGCAGGTGTGTGATGGACAGTGAAGTGGCTGGAAAGACCAAAATC 300

OY 174 AsnSerSerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAla
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 AACAGCTCACCCCTCTCGCTACAGCCGCCAGATTGGGGAATTAACTATCAGGGCT 360

OY 194 GlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGlyGlyValAlaValTyrLeuIys
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GGGCTTACTACTGTACTGTGAGGTGCACTTGTATGAGGAAAGGCTGTCTACTGAG 420

OY 214 LeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGlnGluGluPheSerAlaThr
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 CTGGACTTGTGTGTAACAGGTGTGTGTGGCTCGCTGCGTGGAAATTTCTCAGCCACA 480

OY 234 AlaAlaSerSerLeuGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuAlaLeu
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 GCACCAACCTCTCTGCGGCCGCCAGCTCGTTGTGCCAGTGTCTGGGCTGTGGCGCTG 540

OY 254 ArgProGlySerSerLeuArgIleArgThrLeuProTTPAlaHisLeuAlaAlaPro 273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 541 CGGCCAGGGCTGTCTCTGCGATCGCACCTCCCTGGCTCATTAAGCTGCCCCC 600

OY 274 PheLeuThrTyrPheGlyLeuPheGlnValHis 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 601 TTCTTAACCTTGTGACCTTTTCAAGTTTCAAC 633

RESULT 12
AAV24140
ID AAV24140 standard; DNA; 2856 BP.
XX
AC AAV24140:
XX
DT 28-SEP-1998 (first entry)
XX
DE Homo sapiens BARD1 exon 1.

```

```

XX KW BARD1: ring protein; BRCA1: breast cancer; risk; diagnosis; exon; ss.
XX OS Homo sapiens.
XX FH Key
XX FT exon
XX FT /tag= a
XX FT /number= 1
XX PN W09812327-A2.
XX PD 26-MAR-1998.
XX PF 19-SEP-1997; 97WO-US16842.
XX PR 04-APR-1997; 97US-0042985.
XX PR 20-SEP-1996; 96US-0025296.
XX PR 03-APR-1997; 97US-0042611.
XX PA (TEXA ) UNIV TEXAS SYSTEM.
XX PI Baer R, Bowcock AM;
XX DR WPI: 1998-230317/20.
XX PT DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which
XX PT as breast cancer antigen, BRCA1, binding proteins are useful to
XX PT identify patient having or at risk of developing cancer
XX PS Claim 25; Page 307-308; 348pp; English.
XX
CC The sequence is that of containing a BARD1 exon which can be used in the
CC preparation of the recombinant breast cancer antigen, BRCA1, binding for
CC proteins BARD1, B123, BE2, BE14, BE31 or BE445, or a composition for the
CC detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid
CC sequence, specifically a wild type BARD1 composition for the detection
CC or purification of BRCA1, useful to identify a patient having, or at
CC risk of developing cancer. BARD1 can be used in the preparation of an
CC anti-BARD1 antibody, and in the detection and purification of a BRCA1
CC protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the
CC identification of a binding protein agonist or antagonist that alters
CC the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCA1 or the
CC biological activity of the BRCA1-BARD1, B123, BE2, BE14, BE31 or BE445
CC complex. The antibodies can be used to detect BARD1, B123, BE2, BE14,
CC BE31 or BE445, a specific anti-BARD1 antibody can be used to identify
CC a patient having or at risk of developing cancer.
XX
SO Sequence 2856 BP; 616 A; 778 C; 786 G; 674 T; 2 other;
XX
Alignment Scores:
Pred. No.: 4.74 Length: 2856
Score: 127.00 Matches: 84
Percent Similarity: 36.00% Conservative: 33
Best Local Similarity: 25.85% Mismatches: 122
Query Match: 8.80% Indels: 86
DB: Gaps: 13

US-09-245-198A-4 (1-284) x AAV24140 (1-2856)
OY 14 ProLeuProArgSerLeuGlySerArgAspGlyGlyAlaVal----- 27
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1788 CCCACTCGGCCACGCGGCTGTGCGGCCCGCCCTGTGCTCGGAGTCCCTATT 1847

OY 28 -----ArgGlnAlaGlnPro-----ProAlaProMetaAla 38
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1848 TGGAGCATTTGCGGCGCGGCTGTGCGGCCCGCCCGGACACCGCCCTGTGCGGCC 1907

OY 39 ArgArgSerGlnArgArgArgArgGlyArgGlyGluProGlyThrAlaLeuLeuValPro 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1908 CGCGCTCCCAAGACGCGGAGAGCTTGGCGGTTTTCAGTGCCTGACCTCTCCCT 1967

OY 59 LeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValValSer 78

```





QY 105 LeuAsnProGlnThrGluSerGlnAspProAlaProPheLeuAsnArgLeuValAr 125  
Db 551 GGGGCGCTCCCAAGCGAGTCACACGAGATCCGGCCCT-----590  
QY 125 gProArGArgSerAlaProGlyArgGlyThrArgAlaArgAlaIleAlaIleH 145  
Db 591 -----CGAGCGAGCGAGGAGGCC-GGGACAGCGACGACCACTACCGACCA 642  
QY 145 sTyrgIuValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly-ThrV 165  
Db 643 AGAGGCGCTACACGCTCACTTTTCCCGCAGGT-----GCCGCGCTCAGACACGCG 696  
QY 165 Alser-----GlyTrpGluGluAlaArgIleAsnSerSers 177  
Db 697 TGAATATAACTGTTAGCCAAAGCACAGTGGCAGGACCGGTGTCTCATCATCT 756  
QY 177 exProLeuArgTyrrAsnArgGln 184  
Db 757 CCCCTCTGAAGACAGGAGAACAG 779  
RESULT 15  
AAD25519/c  
ID AAD25519 standard; DNA; 154746 BP.  
XX AAD25519;  
AC AAD25519;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human herpesvirus 2 complete DNA genome.  
XX  
KW Human herpesvirus 2; cytostatic; cancer; immunosuppressive; virucide;  
KW antibacterial; fungicide; protozoacide; antineoplastic; antiinflammatory;  
KW antiparasitic; rheumatoid arthritis; neuroprotective; multiple sclerosis;  
KW immune response; vasotropic; vaccine; gene therapy; autoimmune disease;  
KW vasculitis; ds.  
XX  
OS Human herpesvirus 2.  
XX  
PN WO200176643-A1.  
XX  
PD 18-OCT-2001.  
XX  
PF 06-APR-2001; 2001WO-US11372.  
XX  
PR 07-APR-2000; 2000US-195680P.  
XX  
PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX  
PI Orson FM, Kinsey BM, Bhogal BS;  
XX  
DR MPI; 2002-066308/09.  
XX  
PT Composition for oral delivery of vaccines, comprises expression vector  
PT containing antigenic genomic sequence, bound to aggregated  
PT protein-polycationic polymer conjugate or suspension -  
XX  
PS Disclosure: Page 90-132; 145pp; English.  
XX  
CC The invention relates to a composition comprising an expression vector  
CC bound to an aggregated protein-polycationic polymer conjugate or  
CC suspension. The expression vector contains a promoter polynucleotide  
CC sequence operatively linked to a polynucleotide sequence encoding an  
CC antigen which is a fragment of a gene or genome associated with an  
CC infectious disease, cancer and autoimmune disease such as rheumatoid  
CC arthritis, vasculitis, and multiple sclerosis, pathogenic genomes  
CC consisting of bacterium, fungus, protozoa and virus such as human  
CC immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C  
CC virus (HCV), influenza and respiratory syncytial virus (RSV), and  
CC optionally comprising a nucleotide sequence encoding a cytokine (or a  
CC cytokine expression vector), is useful for inducing an immune response  
CC (systemic and/or mucosal) in an organism. The cytokine expression vector  
CC contains a sequence for granulocyte macrophage colony stimulating factor  
CC (GM-CSF) or Interleukin-12 (IL-12). The polynucleotide sequences encoding

CC the antigen and the cytokine are under transcriptional control of same or  
CC different promoter polynucleotide sequences. The expression vector, as a  
CC DNA vaccine is useful for treating a condition in an organism. The  
CC present sequence is human herpesvirus 2 complete DNA genome related  
CC to the invention.  
XX  
SQ Sequence 154746 BP; 23003 A; 54218 C; 54701 G; 22824 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 621 Length: 154746  
Score: 124.50 Matches: 84  
Percent Similarity: 33.948 Conservative: 28  
Best Local Similarity: 23.458 Mismatches: 116  
Query Match: 8.62% Indels: 102  
DB: Gaps: 13  
US-09-245-198a-4 (1-284) x AAD25519 (1-154746)  
QY 10 AlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaValArgGln 29  
Db 32182 GCGCGCTGCGGCCGCCACGAGCTGTCGCCGCGGCCACAGCGCGCTGCTGCGCTT 32123  
QY 30 AlaGln-----ProProAlaProMetaIleAlaArgArgSerGlnArg 43  
Db 32122 TCCCGCGCTCGAGGAGCCCTGCGCCACCGCGCCCTGCTGCTGCTTGAATCGC 32063  
QY 44 ArgArgGly-----ArgArgGlyGluProGlyThrAlaLeuLeuValProLeuAlaLeu 61  
Db 32062 GCGCGAGGTGGCTTCTGCTCGCGGCAATCGCATCGCTGCTGCGGCGCCATC 32003  
QY 62 GlyLeuGlyLeuAla-----LeuAlaCysLeuGlyLeuLeu--- 74  
Db 32002 TATCCGCGCGCGCTCCAAAGCGCGCCGCAATCTCCACGCGCTGCTACACATTGG 31943  
QY 75 -----AlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAla 91  
Db 31942 CCCCCACAGAGCGAGGTGGAGGAGTGACGCGCGCTT---CCCGTCTCTGGA 31886  
QY 92 GlnGluGluLeuValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGlu 111  
Db 31885 GAACTGATGCGCGCGCTCGAGGGAGCGC----- 31856  
QY 112 GluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaPro 131  
Db 31855 -----CCCGACGCGTTCCTTTCACACCGCGCTACCGCTGCGCTGCGCGCA 31808  
QY 132 LysGlyArgGlyThrArgArgAlaIleAlaIle----- 144  
Db 31807 TCTGGCGCGCCAGGCGCGTGGGGGCGCGCGCGCTGCTCCCTCAGCAGCAGAT 31748  
QY 145 -----HisTyrgIu 147  
Db 31747 CCGCGCGCGCTTCGCGAGCTCGACGCCACACTACGCTTGACTACTACAGCACGCGG 31688  
QY 148 ValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerG 167  
Db 31687 AGACAGCGCTCGCTCAACACAGCTCAATGCAAGTGAGTGGAGCGTGCATATGG 31628  
QY 167 yTrpGluAlaIleArgIleAsn-----SerSerSerPro----- 178  
Db 31627 ACCGAGACAGATAAATGTGCTTCATGAGGCGACCCCTCCACTGCCCCACACAGG 31568  
QY 179 -----LeuArgTyrrAsnArgI 184  
Db 31567 CTGCGAGCATACTCTCCGAGAGAGAGCTACGGGTACCTGCTGCTGGTGGTCAATGCGG 31508  
QY 184 nIleGlyGluPheIleValIleThrArgAlaGlyLeuTyrrTyrrLeuTyrrCysGlnValHisPh 204  
Db 31507 CCTCTGGGCGCTGCTGCTTTCGAGGAGGCTTT-----GCGT 31469  
QY 204 eaSpGluGlyValValValTyrrLeuTyrrLeuAspLeuValAspGlyValLeuAlaIle 224  
Db 31468 TACCATTACACCGCGCGTCACTCAAGCTCGCAGACCCCGTGGCGCTGCGGACT 31409





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 07:42:57 ; Search time 15.0648 seconds  
(without alignments)  
781.906 Million cell updates/sec

Title: US-09-245-198a-4  
Perfect score: 1444  
Sequence: 1 MSLLDFEISARRLPLPSRLG.....FWAHLKAPFLTYGFLFOVH 284

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1268	87.8	249	1	TN12_HUMAN
2	1020	70.6	225	1	TN12_MOUSE
3	109.5	7.6	272	1	TNFS_CHICK
4	108.5	7.5	260	1	TNFS_CANFA
5	106.5	7.4	254	1	TNFS2_HUMAN
6	97	6.7	441	1	CG322_ANTMA
7	95.5	6.6	201	1	TNFB_MACEU
8	94	6.5	280	1	TNFB_MACMU
9	93.5	6.5	690	1	RHO_MICLU
10	93	6.4	310	1	V497_MYCTU
11	92	6.4	240	1	TN14_HUMAN
12	92	6.4	280	1	TNPF_CERTO
13	92	6.4	902	1	NRC4_HUMAN
14	90.5	6.3	760	1	MLH1_MOUSE
15	90.5	6.3	814	1	CADF_HUMAN
16	90	6.2	707	1	JTPI_MOUSE
17	89	6.2	280	1	MDCB_KLEPN
18	89	6.2	316	1	TN11_MOUSE
19	88.5	6.1	278	1	TNPF_RAT
20	88.5	6.1	281	1	TNFB_HUMAN
21	87	6.0	197	1	TNFB_RABIT
22	87	6.0	204	1	TNFB_BOVIN
23	87	6.0	291	1	TN10_MOUSE
24	86.5	6.0	250	1	TNPF_MOUSE
25	86	6.0	139	1	YOPE_BACSU
26	86	6.0	205	1	TNFB_MARMO
27	86	6.0	241	1	TN13_MOUSE
28	85.5	5.9	933	1	VGIB_HSV1
29	85.5	5.9	928	1	VGIB_HSVBP
30	85.5	5.9	932	1	VGIB_HSVBC
31	85	5.9	372	1	LKXB_MOUSE
32	85	5.9	379	1	LKXB_MOUSE
33	84.5	5.9	401	1	AROC_MYCTU

34	84	5.8	284	1	HXL1_HUMAN
35	84	5.8	310	1	TNFC_MARMO
36	84	5.8	575	1	MIS_PIG
37	84	5.8	825	1	ICP0_HSV2H
38	83.5	5.8	416	1	RAGE_BOVIN
39	83.5	5.8	505	1	TUB_MOUSE
40	83.5	5.8	545	1	RTN2_HUMAN
41	83.5	5.8	885	1	VGIB_HSV2S
42	83	5.7	228	1	BIOD_PSEAE
43	83	5.7	653	1	APPI_MOUSE
44	83	5.7	936	1	FHL1_YEAST
45	83	5.7	1217	1	AP4_MOUSE

## ALIGNMENTS

RESULT 1	TN12_HUMAN	STANDARD:	PRT: 249 AA.
AC	043508: ORMU27:		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).		
GN	TNFSF12 OR APO3L OR DR3LG.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	SEQUENCE FROM N.A., AND N-TERMINUS OF SOLUBLE FORM.		
RC	TISSUE=Tonsil, and fetal liver;		
RX	MEDLINE=98070415; PubMed=9405449;		
RA	Chicpeportliche Y., Bourdon P.R., Xu H., Hsu Y.-W., Scott H.,		
RT	Hession C., Garcia I., Browning J.L.;		
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."		
RL	J. Biol. Chem. 272:32401-32410(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Fetal kidney;		
RX	MEDLINE=98228355; PubMed=9560343;		
RA	Martens S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,		
RT	Ashkenazi A.;		
RT	"Identification of a ligand for the death-domain-containing receptor APO3."		
RL	Curr. Biol. 8:525-528(1998).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Tonsil;		
RA	Straussberg R.;		
RT	Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	FUNCTION.		
RC	Pubmed=10085077;		
RA	Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;		
RT	"TWEAK Induces angiogenesis and proliferation of endothelial cells."		
RL	J. Biol. Chem. 274:8455-8459(1999).		
CC	-I- FUNCTION: Binds to FN14 and possibly also to TNFSF12/APO3. Weak inducer of apoptosis in some cell types. Mediates NF-kappaB activation. May promote angiogenesis and the proliferation of endothelial cells.		
CC	-I- SUBUNIT: HOMOTRIMER (POTENTIAL).		
CC	-I- SUBCELLULAR LOCATION: Type II membrane protein and secreted.		
CC	-I- TISSUE SPECIFICITY: Highly expressed in adult heart, pancreas, skeletal muscle, brain, colon, small intestine, lung, ovary, prostate, spleen, lymph node, appendix and peripheral blood lymphocytes. Low expression in kidney, testis, liver, placenta, thymus and bone marrow. Also detected in fetal kidney, liver, lung and brain.		
CC	-I- PTM: The soluble form derives from the membrane form		

DS  
SH563

```
CC      by proteolytic processing.
CC      -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC      -1- CAUTION: Ref.3 sequence differs from that shown due to a
CC      frameshift in position 125.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AF030099; AAC51923.1; -
CC      EMBL: AF055872; AAC39724.1; -
CC      EMBL: BC019047; AAH19047.1; ALT_FRAME.
CC      Genew: HGNC:11927; TNFSF12.
CC      MIM: 602695; -
CC      InterPro: IPR000478; TNF_family.
CC      Pfam: PF00229; TNF; 1.
CC      SMART: SM00207; TNF; 1.
CC      PROSITE: PS00251; TNF_1; FALSE_NEG.
CC      PROSITE: PS50049; TNF_2; 1.
CC      CycloLine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
CC      CHAIN 1 249
CC      CHAIN 94 249
CC      FT      DOMAIN 1 21
CC      FT      TRANSMEM 22 42
CC      FT      DOMAIN 43 249
CC      FT      SITE 139 139
CC      FT      CARBOHYD 139 139
CC      SQ      SEQUENCE 249 AA: 27216 MW: E660843361C28B8A CRC64:
      Query Match 87.8%; Score 1268; DB 1: Length 249;
      Best local Similarity 100.0%; Pred. No. 2e-93;
      Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARSRORRRGRGPGTALVPLALGLALACGLLAVLSGRASISAOEPQOEL 95
DB 1 MAARSRORRRGRGPGTALVPLALGLALACGLLAVLSGRASISAOEPQOEL 60
QY 96 VAEEDDPSELNPQTEESODPAPELNLRLVPRRSAPKGRKTRARRAIAAYEHVPRGOD 155
DB 61 VAEEDDPSELNPQTEESODPAPELNLRLVPRRSAPKGRKTRARRAIAAYEHVPRGOD 120
QY 156 GAQAGVGVSGWEARINSSPLRNROIGETVTRAGLYLYCOVHEDEGKAVYIKLD 215
DB 121 GAQAGVGVSGWEARINSSPLRNROIGETVTRAGLYLYCOVHEDEGKAVYIKLD 180
QY 216 LAVDGVIALRCLEEFSAATASSLGPOLRCQVSGLLALRPGSSLRTRTPMALKAAPFL 275
DB 181 LAVDGVIALRCLEEFSAATASSLGPOLRCQVSGLLALRPGSSLRTRTPMALKAAPFL 240
QY 276 TYFGFLPOVH 284
DB 241 TYFGFLPOVH 249

RESULT 2
TN12_MOUSE STANDARD: PRT: 225 AA.
AC 054907: 09GTP2:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak
GN TNFSF12.
OS Mus musculus (Mouse).
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
```

```
OC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal macrophage;
RX MEDLINE=98070415; PubMed=9405449;
RA Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RT Hesson C., Garcia I., Browning J.L.;
RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that
RT weakly induces apoptosis."
RL J. Biol. Chem. 272:32401-32410(1997).
[2]
RN SEQUENCE OF 83-225 FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=21085660; PubMed=11217651;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Binds to FNL4 and possibly also to TNFRSF12/Apo3. Weak
CC inducer of apoptosis in some cell types. Promotes angiogenesis and
CC the proliferation of endothelial cells. Mediates NF-kappaB
CC activation (By similarity).
CC -1- SUBUNIT: Homotrimer (Potential).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- PTM: The soluble form is produced from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF030100; AAC53517.1; -
CC EMBL: AK020909; BAB32249.1; -
CC MGD: MGI:1196259; Tnfsf12.
CC InterPro: IPR000478; TNF_family.
CC Pfam: PF00229; TNF; 1.
CC SMART: SM00207; TNF; 1.
CC PROSITE: PS00251; TNF_1; FALSE_NEG.
CC PROSITE: PS50049; TNF_2; 1.
CC CycloLine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
CC NON_TER 1 225
CC CHAIN <1 225
CC CHAIN 70 225
CC CHAIN 22 225
CC TRANSMEM <1 21
CC FT      DOMAIN 22 225
CC FT      SITE 69 70
      EXTRACELLULAR (POTENTIAL).
      CLEAVAGE (BY SIMILARITY).
```

[illegible]

```

DR PROSITE PS00251; TNE_1; 1.
DR PROSITE PS50049; TNE_2; 1.
KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 272
FT CHAIN 111 272
FT DOMAIN 1 23
FT TRANSMEM 24 44
FT SITE 45 272
FT DISULFD 110 111
FT CARBOHD 190 229
FT CARBOHD 124 124
FT CARBOHD 146 146
FT CARBOHD 251 251
SQ SEQUENCE 272 AA; 30862 MM; 5409F24ABE53CDD7 CRC64;
Query Match 7.6%; Score 109.5; DB 1; Length 272;
Best Local Similarity 26.4%; Pred. NO. 0.095; Mismatches 58; Indels 33; Gaps 7;
Matches 42; Conservative 26;
OY 139 RRAIAHVEVHPRPQDAGVDTVSQM--EARINSSPLRYNRQIGFIYTRAGLY 197
Db 133 RQPIATHLA-----GVKSNTRYRLVKMKMTTSYAPRSSLSIYHE--GKLKVEAGLY 182
OY 198 LYCYQVHFDEGKA-----VYLKIDLLVGVGLARLCIEFSATVAASSLGPQLRCQV-- 247
Db 183 IYSQVSFCTKAASAPFTLYLYLYLPMEDRL--WKGDFTSTSYA-----LCERQS 233
OY 248 ---SGLLALRESSLRIRLTPRAHLKAPFLYFEELEQV 283
Db 234 IREGVFELRQGDMEVNVDTSTAVNVPNGTYFGMFKL 272
RESULT 4
TNE5_CANFA
ID TNE5_CANFA STANDARD: PRT; 260 AA.
AC 097626;
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand).
GN TNFSF5 OR CD40LG OR CD40L.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Hosie M.H., Willett B.J.;
RT "Adjuvant properties of canine CD40L.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
CC proliferation in the absence of co-stimulus as well as IgE
CC production in the presence of IL-4. Involved in immunoglobulin
CC class switching (By similarity).
CC -I- SUBUNIT: HOMODIMER (By similarity).
CC -I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM (By similarity).
CC -I- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF086711; AAD04375.1; -

```



```

RL EMOB J. 13:616-624(1994).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G2/M
CC (MITOSIS) TRANSITION. G2/M CYCLINS ACCUMULATE STEADILY DURING G2
CC AND ARE ABRUPTLY DESTROYED AT MITOSIS.
CC -1- SUBUNIT: INTERACTS WITH THE CDC2 AND CK2 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES STEADILY DURING G2 AND IS
CC ABRUPTLY DESTROYED AT MITOSIS.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X76123; CA53729.1; -.
DR PIR: S41710; S41710.
DR HSSP: P30274; IYIN.
DR InterPro: IPR004366; Cyclin.
DR InterPro: IPR004367; Cyclin.Cterm.
DR Pfam: PF00134; cyclin_1.
DR SMART: PF02984; cyclin_C_1.
DR SMART: SM00385; CYCLIN_C_2.
DR PROSITE: PS00292; CYCLIN_C_1.
DR CYCLIN: Cell cycle; cell division; Mitosis.
DR PROSITE: 441 AA; 49205 MW; E6E4C037C98880A7 CRC64;
SQ
Query Match 6.7%; Score 97; DB 1; Length 441;
Best Local Similarity 24.0%; Pred. No. 1.6;
Matches 70; Conservative 44; Mismatches 104; Indels 74; Gaps 15;
QY 25 GAVBOAPPPAPMAARRSQR-----RGRGEPGTALLVPLALGLALACGL 73
DB 19 GAIKQK-----MAYEKKRRALGIDGNVTVRGVGGKALPOVSRITRGF-----CAQLI 69
QY 74 -----LAVVSLGSRASLS-----AQEPQAEELVAEDDQPSLEINPQTESQ 114
DB 70 ANAEAAEAENKNSLAVNAKADGALPIKRAVARVPQKKYKSPQEIETISPTERRK 129
QY 115 DPAFLNLVLRPRS-----APKGRKTARARALAAHYEVRPGQ-----DGAQAGVDCV 165
DB 130 --APLEKEITGEKSLKKAAPLTJSTLTARSKASV-VRTPKKEIYVIDADAVNNDLAV 186
QY 166 SGWEE-----ARINSSPLRY---NRQIGFIVTRAGLYLYCOVHFDP---EGKAVYL 212
DB 187 VEYEDMYKKFKSAENDSRPHDYMSQPEINEMK--RAILLIDMLVQVHYKFELESFETLYL 244
QY 213 KLIDLIVGVALRC-----LEEPSATASSLGPQLRLCOVSGGLALRPGS 257
DB 245 TIN-IVDRYVLASKTTSRRELQLLGWSSMLIAKYEIWAPEVNDLVICISDGS 295
RESULT 7
TNFB_MACEU STANDARD: PRT: 201 AA.
AC Q9XT48;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lymphotoxin-alpha precursor (LTA-alpha) (TNF-beta) (Tumor necrosis
DE factor ligand superfamily member 1).
GN LTA OR TNFSF1 OR TNFB.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20284948; PubMed=10826697;

```

```

RA Harrison G.A., Deane E.M.;
RT "cDNA cloning of lymphotoxin alpha (LT-alpha) from a marsupial,
RT Macropus eugenii."
RL DNA Seq. 10:399-403(2000).
CC -1- FUNCTION: Cytokine that in its homotrimeric form binds to
CC TNFRSF1A/TNFR1, TNFRSF1B/TNFR and TNFRSF14/HVEM. In its
CC heterotrimeric form with LTB binds to TNFRSF3/LTBR. Lymphotoxin is
CC produced by lymphocytes and cytotoxic for a wide range of tumor
CC cells in vitro and in vivo.
CC -1- SUBUNIT: Homotrimer, and heterotrimer of either two LTB and one
CC LTA subunits or (less prevalent) two LTA and one LTB subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (homotrimer) and type II membrane
CC protein (heterotrimers) (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF119336; AAD41773.1; -.
DR HSSP: P01374; TNF.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00228; TNF_1.
DR PRINTS: PR01234; TNECROSISCT.
DR PRODOM: PD002012; TNF_abc_1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR CYTOKINE: Glycoprotein; Cytokine; Signal.
KT SIGNAL 1 27
FT CHAIN 28 201
FT CARBOHYD 93 93
SQ SEQUENCE 201 AA; 21536 MW; 8C4C371C85091627 CRC64;
Query Match 6.6%; Score 95.5; DB 1; Length 201;
Best Local Similarity 23.2%; Pred. No. 0.86;
Matches 44; Conservative 27; Mismatches 84; Indels 35; Gaps 9;
QY 107 NPQTESODPAFLRLVLRPRSAPKGRTRARAI--AAHYEVRPGQDGAQAGVCGT 164
DB 30 NPDNSHSSPAD-----PQTAQLSOKSLKREPLKRAHL-----VGDPSVQDSIH-- 75
QY 165 VSGWEARINSSP-LRYNRQI--GEPIVTRAGLYLYCOVHFDEGKA-----VYLKL 214
DB 76 ---W---RANIDHARLIRGFLSNNSLLVPTSGLVFVSQVFGASGASCEITPLLYLSH 129
QY 215 DLVDG---VLALRCLEEPSATASSLGPQLRLCOVSGGLALRPGSSLRIRLTPAHLKA 271
DB 130 EYLLFSSRYQVHVPLLSAQKSVCSGTOGPMWRSVYQGAVFLLTQGDRLSTYDGVSHLQ 189
QY 272 APFLTYFGLF 281
DB 190 SPSSVYFGAF 199
RESULT 8
TNFB_MACMU STANDARD: PRT: 280 AA.
AC Q9MTL6; Q9BDM5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)
DE (CD95L protein).
GN TNFSF6 OR FASL OR CD95L.
OS Macaca mulatta (Rhesus macaque),
OS Macaca fascicularis (Crested eating macaque) (Cynomolgus monkey), and
OS Macaca nemestrina (Pig-tailed macaque).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 NCBI\_TaxID=9544, 9541, 9545;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.mullata; TISSUE=Lymphocytes;  
 RX MEDLINE=21383618; PubMed=1491535;  
 RA Villingger F., Bostlik P., Mayne A.E., King C.L., Genain C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RT "Cloning, sequencing, and homology analysis of nonhuman primate  
 RT Fas/FasLigand and co-stimulatory molecules.";  
 RL Immunogenetics 53:315-328(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.mullata, M.fascicularis, and M.nemestrina;  
 RA Kiril'y, Inoue T., Yoshino K.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to TNFSF6/FAS, a receptor that  
 CC transduces the apoptotic signal into cells. May be involved in  
 CC cytotoxic T cell mediated apoptosis and in T cell development.  
 CC TNFSF6/FAS-mediated apoptosis may have a role in the induction of  
 CC peripheral tolerance. In the antigen-stimulated suicide of mature  
 CC T cells, or both. Binding to the decoy receptor TNFRSF6B/DCR3  
 CC modulates its effects (by similarity).  
 CC -1- SUBUNIT: Homotrimer (Potential).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (by  
 CC similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF344856; AAK37539.1; -;  
 DR EMBL: AB035138; BAA90294.1; -;  
 DR EMBL: AB035139; BAA90295.1; -;  
 DR EMBL: AB035140; BAA90296.1; -;  
 DR HSSP: P01375; 4TSV.  
 DR InterPro: IPR003636; TNF-abc.  
 DR InterPro: IPR000478; TNF-family.  
 DR Pfam: PFO0229; TNF\_1.  
 DR PRINTS: PR01234; TNECROSISFCT.  
 DR ProDom: PD002012; TNF-abc; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 FT Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 280  
 FT CHAIN 129 280  
 FT DOMAIN 1 80  
 FT TRANSMEM 81 101  
 FT DOMAIN 102 280  
 FT SITE 4 69  
 FT DOMAIN 45 64  
 FT SITE 128 129  
 FT DISULFID 201 232  
 FT CARBOHYD 183 183  
 FT CARBOHYD 249 249  
 FT CARBOHYD 259 259  
 FT CONFLICT 60 60  
 FT CONFLICT 31367 MW; F0B284D61A132EB4 CR664;  
 SO SEQUENCE 280 AA; 31367 MW; 6.5%; Score 94; DB 1; Length 280;

Best Local Similarity 20.8%; Pred. No. 1.6;  
 Matches 62; Conservative 45; Mismatches 117; Indels 74; Gaps 14;  
 QY 13 LPLPRISGSRDGAIVQAQAPAPMAARRSQRGRGEPPTALLVPLAL----- 61  
 DB 30 LPPPTSVPRRRGRRPPPPPPPPPP-----PPPPSPPLPPLPKKGNHST 78  
 QY 62 GLGLALACGLLALLAVVSLGSRASLSAEPAGEELVAEEDDPSELNPQTEESDPAFLN 121  
 DB 79 GLCLLWFEVFLVALVGLG- LGMFQLFHQKEL-----ALRESTSQKRTASLEK 128  
 QY 122 RLVRPRRSAPKGRKTRARRAIAAHYEHVPRPGDGAQAGVDGVSGWEA- RINSSPLR 180  
 DB 129 QIGHP---SPPEKKEDQRK--VAHLTKPMSRSMPL-----WEDTYGIVLLSGVK 174  
 QY 181 YNIOIGEFIVTRAGLYLVYQVHFDEGKA-----VYLKID-----LVDGYLALR 225  
 DB 175 YKR--GGLVINEGTLGVSKVYF-RQOSTNLP LSHKRYMRSKYPQDLVMEGKMST 231  
 QY 226 CLEEFSTAAASLGLPOLRLCQVSGLLALRPSSLRITLFWAHKAPFLTYGRLPOV 283  
 DB 232 CTTGQWMAHSSYLGAVENLTSADHLV-----VNVSELVLVFEESQ--TFGLYKL 280  
 RESULT 9  
 RHO\_MICLU STANDARD; PRT; 690 AA.  
 AC P52154;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transcription termination factor rho.  
 GN RHO.  
 OS Micrococcus luteus (Micrococcus lysodeikticus).  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.  
 OX NCBI\_TaxID=1270;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5 AND 289-297.  
 RC STRAIN-EM;  
 RX MEDLINE=96132802; PubMed=8557681;  
 RA Nowatzke W.L., Richardson J.P.;  
 RT "Characterization of an unusual Rho factor from the high G + C gram-  
 RT positive bacterium Micrococcus luteus.";  
 RL J. Biol. Chem. 271:742-747(1996).  
 RN [2]  
 RP SEQUENCE OF 205-690 FROM N.A.  
 RC STRAIN-EM;  
 RX MEDLINE=94327472; PubMed=8051015;  
 RA Opperman T., Richardson J.P.;  
 RT "Phylogenetic analysis of sequences from diverse bacteria with  
 RT homology to the Escherichia coli rho gene.";  
 RL J. Bacteriol. 176:5033-5043(1994).  
 RN [3]  
 RP REVISION TO 500.  
 RC STRAIN-EM;  
 RA Nowatzke W.L.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM  
 CC THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF RHO'S  
 CC RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE RNA FROM THE  
 CC DNA TEMPLATE. RNA-DEPENDENT NPASE WHICH UTILIZES ALL FOUR  
 CC RIBONUCLEOSIDE TRIPHOSPHATES AS WELL AS DAPF AS SUBSTRATES, BUT  
 CC HAS A SIGNIFICANT LOWER ACTIVITY WITH CTP.  
 CC -1- SUBUNIT: HOMOHXAMER (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>





RT "LIGHT, a new member of the TNF superfamily, and lymphotoxin alpha are  
RT ligands for herpesvirus entry mediator.";  
RL Immunity 8:21-30(1998).  
[2]  
RN SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.  
RC TISSUE-LIVER;  
RX MEDLINE-98438532; PubMed-9765287;  
RA Harrop J.A., McDonnell P.C., Brigham-Burke M., Lyn S.D., Minton J.,  
RA Tan K.B., Dede K., Spemanato J., Silverman C., Hensley P.,  
RA DiPonzio R., Emery J.G., Deen K., Eichman C., Chabot-Fletcher M.,  
RA Truneh A., Young P.R.;  
RT "Herpesvirus entry mediator ligand (HVEM-L), a novel ligand for  
RT HVEM/RT2, stimulates proliferation of T cells and inhibits HT29 cell  
RT growth.";  
RL J. Biol. Chem. 273:27548-27556(1998).  
[3]  
RN SEQUENCE FROM N.A. (ISOFORM 2), AND PROCESSING.  
RX MEDLINE-21528948; PubMed-11673523;  
RA Granger S.W., Butrovich K.D., Houshmand P., Edwards W.R., Ware C.F.;  
RT "Genomic characterization of LIGHT reveals linkage to an immune  
RT response locus on chromosome 19p13.3 and distinct isoforms generated  
RT by alternate splicing or proteolysis.";  
RL J. Immunol. 167:5122-5128(2001).  
[4]  
RN SEQUENCE FROM N.A.  
RC TISSUE-Brain;  
RX Strausberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to the  
CC decay receptor TNFRSF6B modulates its effects. Activates NFkB,  
CC stimulates the proliferation of T cells, and inhibits growth of  
CC the adenocarcinoma HT-29. Acts as a receptor for Herpes simplex  
CC virus.  
CC -1- SUBUNIT: Homotrimer.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted  
CC (isoform 1); Cytoplasmic (isoform 2).  
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/LIGHT  
CC delta-TM; are produced by alternative splicing.  
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT ALSO  
CC FOUND IN THE BRAIN. WEAKLY EXPRESSED IN PERIPHERAL LYMPHOID.  
CC TISSUES AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KIDNEY,  
CC AND NO EXPRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OR  
CC NONHEMATOPOIETIC TUMOR LINES.  
CC -1- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.  
CC -1- PTM: N-glycosylated.  
CC -1- PTM: The soluble form of isoform 1 derives from the membrane  
CC form by proteolytic processing.  
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
CC -1- CAUTION: Ref.4 sequence differs from that shown due to a  
CC frameshift in position 178.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF036581; AAC39563.1; -;  
DR EMBL: AF064090; AAC25169.1; -;  
DR EMBL: AY028261; AAK26160.1; -;  
DR EMBL: BC018058; AAH18058.1; ALT\_FRAME.  
DR HSSP: P01375; 4TSV.  
DR Gene: HGNC:11930; TNFSF14.  
DR MIM: 604520; -;  
DR InterPro: IPR003636; TNF\_abc.  
DR InterPro: IPR000478; TNF\_family.  
DR Pfam: PF00229; TNF\_1.  
DR ProDom: PD002012; TNF\_abc; 1.  
DR SMART: SM00207; TNF; 1.  
DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
DR PROSITE: PS50049; TNF\_2; 1.

KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;  
KW Alternative splicing.  
FT CHAIN 1 240  
FT CHAIN 1 240  
FT CHAIN 783 240  
FT CHAIN 1 240  
FT CHAIN 37 37  
FT CHAIN 38 58  
FT CHAIN 59 240  
FT CHAIN 82 83  
FT CHAIN 154 187  
FT CHAIN 102 102  
FT CHAIN 38 73  
FT CHAIN 120 120  
FT CHAIN 214 214  
FT CHAIN 240 AA; 26351 MW; 490DBF67EL390839 CRC64;  
SO SEQUENCE  
Query Match 6.4%; Score 92; DB 1; Length 240;  
Best Local Similarity 23.7%; Pred. No. 2;  
Matches 44; Conservative 19; Mismatches 57; Indels 66; Gaps 6;  
QY 29 QAOPEPMARSRORRGRGEPGTALVPLALGLALACGLLAVVSGRASLSAQ 88  
DB 16 QTDIPFTRGLGRSHRQSCSVARVGLLL-LMGAGLAAGWGLQLHMRFG----- 66  
QY 89 EPAQEELVAEEDDPSELNPQTEESODPAPFLNRLVPRRSAPGRKTRARRAAHVEY 148  
DB 67 -----EVV-----TLPPGAPAGSEQLQGRSS-----HEV 92  
QY 149 HPRPGDGAQAGVDGYSGWEARINSSPLRYNRQI-----GEFVTRAGLY 196  
DB 93 NPAHLTGANSLSLTG-----SGGPLMETWGLGLAFLRGLSYHDGALVTRAGVY 141  
QY 197 YLYCOV 202  
DB 142 YLYSKV 147  
RESULT 12  
TNF6\_CERTO STANDARD; PRT; 280 AA.  
ID TNF6\_CERTO  
AC Q9BUNI;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)  
DE (CD95L protein).  
GN TNFSF6 OR FASL OR CD95L.  
OS Cercopithecus torquatus alyx (Red-crowned mangabey) (Sooty mangabey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Cercopithecidae; Cercopithecus.  
OX NCBI\_TaxID=9531;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Lymphocytes;  
RX MEDLINE-21383618; PubMed-11491535;  
RA Villinger F., Bostlik P., Mayne A.E., King C.L., Genain C.P.,  
RA Weiss W.R., Ansari A.A.;  
RT "Cloning, sequencing, and homology analysis of nonhuman primate  
RT Fas/Fas-Ligand and co-stimulatory molecules.";  
RL Immunogenetics 53:315-328(2001).  
CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that  
CC transduces the apoptotic signal into cells. May be involved in  
CC cytotoxic T cell mediated apoptosis and in T cell development.  
CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of  
CC peripheral tolerance. In the antigen-stimulated suicide of mature  
CC T cells, or both. Binding to the decay receptor TNFRSF6B/DCR3  
CC modulates its effects (by similarity).  
CC -1- SUBUNIT: Homotrimer (Probable).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
CC similarity).



FT DOMAIN 672 674 NUCLEAR LOCALIZATION SIGNAL.  
 SQ SEQUENCE 902 AA: 95472 MW: E59F15F7647A47C6 CRC64;  
 Query Match  
 Best Local Similarity 25.0%; Score 92; DB 1; Length 902;  
 Matches 48; Conservative 15; Mismatches 73; Indels 56; Gaps 8;  
 QY 14 PLPRSLGSRDGGAVRQAPPAWMAARRS---QRRRGRRGETALLVPLALGLALAC 69  
 DB 236 PPSGCGPEPDSWLLSAPPTPASPRPAPCGKRRYSSSGTPSSA-----SPALSR 286  
 QY 70 LGILLAVSLGSRASLSAOPAEOLVAEDDPSSELNQTESQDPAPPLNLRVPRRS 129  
 DB 287 RG-----SLGEGS-----EPPPPPL-PLARDPGSPGPFYVGAPPAES 325  
 QY 130 AP-KGRKTRARRAIA-----AHYEVHPRPGDGAQAGVD-----GTVSG 167  
 DB 326 IQQKTRRTSSEQVALLPRSEEPASCKLPLGAESVAPPGSKKEVAGMDYLAVPSPLA 385  
 QY 168 WEARINSSSPL 179  
 DB 386 WSKARIGHSP 397  
 RESULT 14  
 MLH1\_MOUSE STANDARD: PRT: 760 AA.  
 ID MLH1\_MOUSE 062454;  
 AC 16-OCT-2001 (rel. 40, Created)  
 DT 16-OCT-2001 (rel. 40, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE DNA mismatch repair protein Mlh1 (Mult. protein homolog 1).  
 GN MLH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Testis;  
 RA Kumaran M., Rao M.R.S.;  
 RT "Cloning of the cDNA of the Mult. homolog, MLH1 from mouse testis.";  
 RL Submitted (Mar-2000) to the EMBL/Genbank/DBJ databases.  
 RP SEQUENCE OF 1-151 FROM N.A.  
 RX MEDLINE=96270514; PubMed=8674118;  
 RA Edelman W., Cohen P.E., Kane M., Lau K., Morrow B., Bennett S.,  
 RA Umar A., Kunkel T., Catoretti G., Chaganti R., Pollard J.W.,  
 RT "Meiotic pachytene arrest in Mlh1-deficient mice.";  
 RL Cell 85:1125-1134(1996).  
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE REPAIR OF MISMATCHES IN DNA.  
 CC -1- SUBUNIT: HETERODIMER OF MLH1 AND PMS2 OR MLH1 AND MLH3 (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: AF250844; AAF64514.1;  
 DR EMBL: U60872; AAC52672.1;  
 DR EMBL: U59881; AAC52672.1; JOINED.  
 DR EMBL: U59883; AAC52672.1; JOINED.  
 DR EMBL: U59884; AAC52672.1; JOINED.  
 DR HSSP: P23367; IBKN.  
 DR MGD: MGI:101938; Mlh1.  
 DR InterPro: IPR003594; ATPbind\_Arase.

DR InterPro: IPR002099; DNA\_mis\_repair.  
 DR Pfam: PF01119; DNA\_mis\_repair; 1.  
 DR Pfam: PF02518; HATPase\_C; 1.  
 DR TIGRPFAM: TIGR00585; mutL; 1.  
 DR PROSITE: PS00058; DNA\_MISMATCH\_REPAIR\_1; 1.  
 KW DNA repair.  
 SQ SEQUENCE 760 AA: 84679 MW: 173C809372A29186 CRC64;  
 Query Match  
 Best Local Similarity 22.7%; Score 90.5; DB 1; Length 760;  
 Matches 70; Conservative 31; Mismatches 126; Indels 81; Gaps 14;  
 QY 13 LPLPRSLGSRDGGAVRQ-----AOPPAWMAARRSQRRRGRRGETALLVPLALGL 64  
 DB 436 LPAPAEAAESENLERESIMETSDAAQKAPTSPPSSRRHR--BDSDVENAENASCKE 493  
 QY 65 LALAC-----LGILLAVSLGSRASLSAOPAEOLVAEDDPSSELNQ--TEEQDRA 117  
 DB 494 MTAACYPRRRRIINLTSLQOEISERCHETURE--ILNHSFVGCVNQWALAQHOTKL 551  
 QY 118 PFLNRLVRRRSAPKGRKTRARRAIAHYEV-----HPRPGDGAQAGVDG 163  
 DB 552 YLLN-----TTKLSEELFYQILLYDRANGVLRSPAPLFDLAMLALDS 596  
 QY 164 TVSGWEARINSSSPLRYNRQIGFTVTRAGLYLYCQVHDEGKAVYLUKLDLYDG--- 220  
 DB 597 PESGWTEDDQPKEGLAEV---IVEFLKKAEMLADYFSEIDE--EGNLIGLPLIDSYVP 652  
 QY 221 -----VLALR-----CLEEFSNTAAS--SLGPQ--LRICQVSGILLALRPGSS 258  
 DB 653 PLEGPLIFILRLATEVNMDEKECFESLSKECAMFYIRKQYILLESTLSGQSDMPGST 712  
 QY 259 LRIRTLPM 266  
 DB 713 SK-----PW 716  
 RESULT 15  
 CADF\_HUMAN STANDARD: PRT: 814 AA.  
 ID CADF\_HUMAN P55291;  
 AC 01-OCT-1996 (rel. 34, Created)  
 DT 01-OCT-1996 (rel. 34, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Muscle-cadherin precursor (M-cadherin) (Cadherin-15) (Cadherin-14).  
 GN CDH15 OR CDH14 OR CDH3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Skeletal muscle;  
 RA MEDLINE=97184182; PubMed=9030594;  
 RA Shibata T., Shimoyama Y., Goton M., Hirohashi S.;  
 RT "Identification of human cadherin-14, a novel neurally specific type  
 RL II cadherin, by protein interaction cloning.";  
 RL J. Biol. Chem. 272:5236-5240(1997).  
 RP SEQUENCE FROM N.A.  
 RC Tissue-Muscle;  
 RA Strausberg R.;  
 RL Submitted (May-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS. CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. M-CADHERIN IS PART OF THE  
 CC MYOGENIC PROGRAM AND MAY PROVIDE A TRIGGER FOR TERMINAL MUSCLE  
 CC DIFFERENTIATION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 31, 2003, 01:56:48 ; Search time 1878.64 Seconds

(without alignments)  
2448.323 Million cell updates/sec

Title: US-09-245-198a-4  
Perfect score: 1444  
Sequence: 1 MSLDFEISARLPLPSIG.....FWAHLKAPFLTYGFLQVH 284

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n\_model -DEV=xlh  
-O=/qgr2.1/USPRO/spool/US09245198/runat\_24032003.163554.25956/app\_query.fasta.1.846  
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09245198 @CGN.1.1.2463 @runat.24032003.163554.25956 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_XMAP -LARGQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estln:\*  
4: em\_estlm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlm:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	1087.5	75.3	731	13	BI871711 603395825
2	1060.5	73.4	728	13	BI870393 603395641
3	979	67.8	777	13	BI819200 603034614
4	962	66.6	834	13	BI766766 603056866
5	934	64.7	828	13	BI596681 603243254
6	913	63.2	918	13	BE577781 602092080
7	909.5	63.0	948	14	BE0707185
8	904	62.6	940	14	BE084231
9	825	57.1	561	10	BE0763237
10	809	56.0	963	14	BE071259
11	765	53.0	1071	14	BE921213
12	753	52.1	951	14	BE074188
13	742.5	51.4	785	13	BE762908
14	738	51.1	587	12	BE686319
15	707	49.0	1033	11	AK020909
16	652	44.9	584	10	AW917574
17	648.5	44.9	698	13	BI906850
18	647	44.8	894	13	BI908274
19	645.5	44.7	474	13	BI965174
20	604	41.8	445	9	AA870722
21	598.5	41.4	471	9	AA221610
22	582.5	40.3	542	12	BE041509
23	582	40.3	650	12	BE040436
24	572.5	39.6	404	12	BE044430
25	562	38.9	360	10	BE654876
26	520.5	36.0	493	10	BE307031
27	506	35.0	538	12	BE821434
28	499	34.6	558	13	BE484863
29	493	34.1	298	14	BE688946
30	476	33.0	441	13	BI967060
31	466	32.3	315	12	BE466521
32	441	30.5	465	9	AT091441
33	440	30.5	581	13	BI738634
34	421	29.2	345	14	RS5379
35	420	29.1	278	10	AA417023
36	403	27.9	1187	14	BE053284
37	395.5	27.4	571	12	BE073881
38	391	27.1	639	10	BE642326
39	388.5	26.9	1926	14	BE906056
40	363	25.1	400	9	AI152313
41	362	25.1	342	9	AA637970
42	353	24.4	910	12	BE110063
43	352	24.4	456	13	BI966235
44	347	24.0	487	10	AW320117
45	320	22.2	533	10	BE628951

## ALIGNMENTS

RESULT 1  
BI871711  
LOCUS 603395825P1 NIH\_MGC\_90 Homo sapiens CDNA clone IMAGE:5405478 5',  
DEFINITION mRNA sequence.  
ACCESSION BI871711 731 bp mRNA  
VERSION BI871711.1 GI:16045386  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 731)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE

JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph. D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLAM12034 row: b column: 07  
High quality sequence stop: 728.  
Location/Qualifiers

FEATURES  
Source  
1. 731  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5405478"  
/clone\_lib="NIH\_MGC\_90"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Liver; Vector: PCMV-SPORE6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 125 a 242 c 228 g 136 t  
ORIGIN

## Alignment Scores:

Pred. No.:	2,28e-93	Length:	731
Score:	1087.50	Matches:	219
Percent Similarity:	99.10%	Conservative:	0
Best Local Similarity:	99.10%	Mismatches:	1
Query Match:	75.31%	Indels:	2
DB:	13	Gaps:	1

US-09-245-198a-4 (1-284) x B1871711 (1-731)

```
OY 65 LeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSer 84
|||||
DB 2 CTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 61

OY 85 LeuSerAla---GlnGluProAlaGlnGlnGluLeuValaGlnGluAspGlnAspPro 103
|||||
DB 62 CTGTCGCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121

OY 104 SerGluLeuAsnProGlnThrGlnGluSerGlnAspProAlaProPheLeuAsnArgLeu 123
|||||
DB 122 TCGGAAGTGAATCCCGACAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 181

OY 124 ValArgProArgArgSerAlaProGlyGlyArgGlyThrArgAlaArgArgAlaIleAla 143
|||||
DB 182 GTTCGGGCTCGCAGAACTGCACTAAAGGCCGGAACACAGGGGCTCGAAGAGCGATCGCA 241

OY 144 AlaHisTyrGlnValHisProArgProGlnGlnAspGlnValaGlnAlaGlyValaAspGly 163
|||||
DB 242 GCCCATTTATAGTTCTTCACGACCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301

OY 164 ThrValSerGlyTyrGlnGlnAlaArgIleAsnSerSerSerProLeuArgTyrAsnArg 183
|||||
DB 302 ACGATGAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361

OY 184 GlnIleGlyGlnPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrGlnValHis 203
|||||
DB 362 CAGATCGGGAGATTATAGTCACCCGGGCTGCTACTACCTGATCTGACAGGTGCAC 421

OY 204 PheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuValaAspGlyValaIleuAla 223
|||||
DB 422 TTTGATGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481

OY 224 LeuArgCysLeuGlnGlnPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg 243
|||||
DB 482 CTGCGCTGGCTGGAGAAATTTCTAGCCACTGCGGCGAGTTCCCTCGGGGCCCGACGCTCGC 541
```

```
OY 244 LeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThr 263
|||||
DB 542 CTGCGCAGAGTGTCTGGGCGCTTGGCCCTCGGCGACAGTGC-TCCCTGGGAGATCGGAC 600

OY 264 LeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283
|||||
DB 601 CTCCTCCCGGGGCCATCTCAAGGCTGGCCCTTCTCTCACTTCTGAGCTTCCAGGTT 660

OY 284 His 284
|||
DB 661 CAC 663
```

RESULT 2  
B1870393  
LOCUS B1870393 728 bp mRNA linear EST 11-OCT-2001  
DEFINITION 603395641F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:5405459 5',  
mRNA sequence.  
ACCESSION B1870393 GI:16044066  
VERSION B1870393  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph. D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLAM12034 row: a column: 12  
High quality sequence stop: 728.  
Location/Qualifiers

FEATURES  
Source

1. 728  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5405459"  
/clone\_lib="NIH\_MGC\_90"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Liver; Vector: PCMV-SPORE6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## BASE COUNT

125 a 240 c 227 g 136 t

## ORIGIN

Alignment Scores:  
Pred. No.: 8.44e-91 Length: 728  
Score: 1060.50 Matches: 218  
Percent Similarity: 98.64% Conservative: 0  
Best Local Similarity: 98.64% Mismatches: 2  
Query Match: 73.44% Indels: 3  
DB: 13 Gaps: 1

US-09-245-198a-4 (1-284) x B1870393 (1-728)

```
OY 65 LeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSer 84
|||||
DB 2 CTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 61

OY 85 LeuSerAla---GlnGluProAlaGlnGlnGluLeuValaGlnGluAspGlnAspPro 103
|||||
DB 62 CTGTCGCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121
```



QY	104	SerGluLeuAspProGluThrGluLeuSerGlnAspProAlaProPheLeuAsnArgLeu	123
Db	122	TGGAACTGAATCCCCAGACAGAAAGAACCCAGGATCTCGCCCTTCTCTGAACCGACTA	181
QY	124	ValArgProArgArgSerAlaProIleuSglYlaArgVsthrArgAlaArgAlaIleAla	143
Db	182	GTTGGCGCTCGAGAAAGTGACACTTAAAGGCCGAAACACGGGGCTCGAAGACGATCGCA	241
QY	144	AlaHisTYRGLuValHisProArgProGluGlnAspGlyAlaGlnAlaGlyValAspGly	163
Db	242	GCCCAATTGTGAAGTTCATCCACAGACTGGACAGGACGGAGCGCAGCGATGTGGACGGG	301
QY	164	ThrValSerGlyTYRGLuGlnAlaArgIleAsnSerSerProLeuArgTYrAsnArg	183
Db	302	ACAGTAGAGTGGCTGGAGAGAACCCAGAAATCAACAGACTCCAGCCCTTGCGCTACAAACCGC	361
QY	184	GlnIleGlyGluPheIleValThrArgAlaGlyLeuTYRTYRLeuTYRLeuTYRCysGlnAlaHis	203
Db	362	CAGATCGGGGAGATTATAGTGCACCCGGGGTGGGCTCTACTACTGTACTGTACGGTGCAC	421
QY	204	PheAspGluGlyLysAlaValTYrLeuLysLeuAspLeuValAspGlyValLeuAla	223
Db	422	TTTGATGAGGGGAAAGCTGTCTACTCTGAAAGCTTGACCTTGCTGGATGTGCTGGCC	481
QY	224	LeuArgCysLeuGlnGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg	243
Db	482	CTGGCTCTCGCTGGAGGAATCTTCAGCCACATGGGCCAGATGCTCTGGGGCCCGAGCTCGCC	541
QY	244	LeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThr	263
Db	542	CTCTCCAGAGTGTGTGGGCTGTTGGCCCTGGGGCCAGGATCC-TCCCTGGAGATCCGCACCC	600
QY	264	LeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTYRPhcGlyLeuPheGlnVal	283
Db	601	CTCCCTCGGGCCATCTCAAGGCTGC-CCCTTCCTCACCTACTTGAGACTCTTCAGAGTT	659
QY	284	His 284	
Db	660	CAC 662	
RESULT 3			
LOCUS	BI819200	777 bp	mRNA linear EST 04-OCT-2001
DEFINITION	603034614P1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175698 5',		
ACCESSION	BI819200		mRNA sequence.
VERSION	BI819200.1	GI:15930750	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 777)		
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strusberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: L14M11437 row: 1 column: 03 High quality sequence stop: 759. Location/Qualifiers 1..777 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5175698"		
FEATURES	source		

Accession	Gene	Protein	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB
US-09-245-198A-4 (1-284)	x B1819200 (1-777)							
OY	97	AlaGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro	116	5.33e-83	777	797.00	100.00%	100.00%
Db	14	GCAGAGGAGGACACAGGACCCGTCGGAGACTGATCCCGACAGAGAAAGCCAGATCCT	73			Matches: 188	Conservative: 0	Mismatches: 0
OY	117	AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyGlyArgGlyThr	136			Indels: 0	Gaps: 0	
Db	74	GGCGCTTTCCTGAACCGACTAGTGTGGCTCGCAAGAGGACCTTAAAGGCGGAAACA	133					
OY	137	ArgAlaArgArgAlaIleAlaAlaHisIstyrgIuValHisProArgProGlyGlnAspGly	156					
Db	134	CGGCTCGAAGAGCATGCGACGCCCATTTATGAGTTCCACGACCTGGACAGGACCGA	193					
OY	157	AlaGlnAlaGlyValAspGlyThrValSerGlyIstyrgIuAlaArgIleAsnSerSer	176					
Db	194	GGCGAGGACGCTGGAGCGGACAGTGAAGTGGCTGGAGAGAAAGCCAGATCAACAGCTCC	253					
OY	177	SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuIst	196					
Db	234	AGCCCTCGCGCTACACCGCCAGATCGGGGAGTTTATGATACCCGGGCTGGCTTAC	313					
OY	197	TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrIleuLysLeuAspLeu	216					
Db	314	TACCTGTACTGTCAAGGTGACACTTATATGAGGAGGAGGCTGTCTACCTGAGCTGATG	373					
OY	217	LeuValAspGlyValIleuAlaLeuArgCysLeuGluGluIlePheSerAlaThrAlaAser	236					
Db	374	CTGGTGTATGGTGCTGTGGCCCTCGGTCGCGGAGAAATTCACAGCCACTCGGCGAGT	433					
OY	237	SerLeuGlyProGlnIleuArgIleuGlyCysGlnValSerGlyLeuAlaLeuArgProGly	256					
Db	434	TCCCTCGGGGCCACTCGCTCGCCAGAGGTGTGGGCTGTGGCCCTGGGCGGACAGG	493					
OY	257	SerSerLeuArgIleArgThrLeuProPheAlaHisLeuLysAlaAlaProPheLeuThr	276					
Db	494	TCCCTCGCTGGGATCGGACACCTCCCTCTGGGCGCATCTCAAGGCTGCCCCCTTCTCAC	553					
OY	277	TyrPheGlyLeuPheGlnValHis	284					
Db	554	TACTTCGGACCTTCCAGGTTTAC	577					

```

SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 834)
AUTHORS     NIH-MGC http://mgi.ncl.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgaabs-remail.nih.gov
             Tissue Procurement: Life Technologies, Inc.
             CDNA Library Preparation: Life Technologies, Inc.
             CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLM11517 row: c column: 18
             High quality sequence stop: 772.
FEATURES
    source
        1..834
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5206217"
            /clone_1lb="NIH-MGC_122"
            /lab_host="DH10B"
            /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
            Site_1: NotI; Site_2: EcoRV (destroyed); RNA source:
            anonymous pool of 24 week female lung, 16 week female
            spleen, and 20-22 week male spleen. Library is oligo-dT
            primed and directionally cloned (EcoRV site is destroyed
            upon cloning). Average insert size 1.4 kb, insert size
            range 1-3 kb. Library is normalized and enriched for
            full-length clones and was constructed by C. Gruber
            (Invitrogen). Research Genetics tracking code 026. Note:
            this is a NIH-MGC Library."
BASE COUNT   154 a      287 c      223 g      170 t
ORIGIN
Alignment Scores:
Pred. No.:      2,46e-81      Length:      834
Score:          962.00        Matches:      192
Percent Similarity: 98.97%      Conservative: 0
Best Local Similarity: 98.97%      Mismatches: 1
Query Match:    66.62%      Indels:      2
DB:             13          Gaps:          0
US-09-245-198a-4 (1-284) x B1766766 (1-834)
OY      92  GlnGluGluLeuValAlaGluGluAspGluAspProSerGluLeuAsnProGlnThrGlu 111
          |||||||
Db      3  CAGAGGAGAGCTGTGTCAGAGAGAGACAGACCCGCGAGACTGATCCCGACAGAGAA 62
OY      112 GluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArgSerAlaPro 131
          |||||||
Db      63  GAAAGCCAGAGATCCCTGCTTCGAAACCGACTAGTTGGCTCCGACAGATGCACCT 122
OY      132 LysGluValArgLysThrArgAlaArgAlaIleAlaIleHisThrGluValHisProArg 151
          |||||||
Db      123  AAAGGCCGGAACACGGGCTCCAGAGAGCATGCGACCCATATGAAAGTTCCACGA 182
OY      152 Pro-GluGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluAla 171
          |||||||
Db      183  CCGTGACAGAGAGAGGAGGAGGAGGTGTGAGAGGAGACAGTAGTGCTGGGAGAGAGC 242
OY      171  AatGgtLeasnSerSerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValTh 191
          |||||||
Db      243  CAGAATCAACAGAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATGGGGAGTTTATAGCAC 302
OY      191  ratGatAglyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGlyLysAlaValTy 211
          |||||||
Db      303  CCGGGCTGGGCTTACTACTGACTGTACAGGTGACACTTTGATGAGGGAGGCTGTCTA 362

```

```

OY      211  rLeuLysLeuAspLeuLeuValAspGlyValIleuAlaIleuArgCysLeuGluPheSe 231
          |||||||
Db      363  CTTGAAGCTGAGACTCTGTGTGATGGTGTGGCCCTGCTGCTGCGAGAAATTTCTC 422
OY      231  ratThrAlaIleSerSerLeuLysProGlnIleuArgLeuCysGlnValSerGlyLeuLe 251
          |||||||
Db      423  AGCCACTGGCGGCGAGTTCCTCGGGCCCGAGCTCCGCTGTGCAAGATGTGGGCTGT 482
OY      251  uAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAl 271
          |||||||
Db      483  GGGCCCTGGCGGCGAGGTCC-TCCCTGGGGATCGCACCCCTCCCTGGGGCCATCTCAAGGC 541
OY      271  aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
          |||||||
Db      542  TGGCCCTTCTCCTCAGCTACTTCTGAGCTTCTCCAGGTTTAC 581
RESULT 5
BI596681      828 bp      mRNA      linear      EST 07-SEP-2001
LOCUS        603243254F1 NIH-MGC_96 Homo sapiens CDNA clone IMAGE:5283892 5',
DEFINITION   mRNA sequence.
ACCESSION   BI596681
VERSION     BI596681.1 GI:15489620
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 828)
AUTHORS     NIH-MGC http://mgi.ncl.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgaabs-remail.nih.gov
             Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
             CDNA Library Preparation: Michael J. Brownstein (NHGRI), ShiraKl
             Toshituki and Piero Carninci (RIKEN)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLM11722 row: k column: 13
             High quality sequence stop: 776.
FEATURES
    source
        1..828
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5283892"
            /clone_1lb="NIH-MGC_96"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /note="Organ: brain; Vector: pBluescript (modified
            pluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
            ); Oligo-dT primed using primer 5'-TTTCTTTTCTTTTCTTTT-3',
            size-selected for average insert size 2.3 kb and
            normalized to R0T 5. This is a primary library enriched
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, In preparation). Library
            constructed by M. Brownstein (NHGRI/NHGRI, National
            Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT   155 a      278 c      223 g      172 t
ORIGIN
Alignment Scores:
Pred. No.:      1.13e-78      Length:      828
Score:          934.00        Matches:      179
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.44%      Mismatches: 0
Query Match:    64.68%      Indels:      0
DB:             13          Gaps:          0
US-09-245-198a-4 (1-284) x BI596681 (1-828)

```

Oy		105	GlutEuaAnPrgInThrGluGlnSerClnsppRoAlaProPhleAAsnArgEuaVal	124
Db		44	GAACTGAATCCCGACAGAAAGAACCAGGACTCTGGCTTCTTAACCGACTAGTT	103
Oy		125	ArgProAAtgAcSserAlaProLysGlyArgrglThrAgaIaaArgAraIleAlaIa	144
Db		104	CGGCCTGCCACAAGTCACCTAAAGGCCGGAAMACAGGGGCTCGAAGACGTGCACACC	163
Oy		145	HlsYrrGluValHlsProArqPrgGlyClnAspgIyaIagIlnalagIyaIaspgIylThr	164
Db		164	CATTATGAAGTTCATCCACGACCTGGACAGGACGCGACGGCAGTGTCGACGGGACA	223
Oy		165	ValSerGlyTPrgUglUAlaAlaArgrlIeAnsSerSerProLeuArGrTrAsmArgIn	184
Db		224	TTCAGTGCCCTGGAGAGAACCCAGAAATCAACAGCTCACGCCCTCTGCGCTACACCGCCAG	283
Oy		185	IleaglUglUpheIleValThrArqIlaaglYleuTyrrTyrlYleuTyrcysGlnValHisphe	204
Db		284	ATCGGGAGATTATATAGTACCCCGGGCTGGGCTCTACTACTGTACTGTACAGTGCACATT	343
Oy		205	AspGluGlyrsAlaValTyrIleuLysLeuAspspleuLeuValAspgIlyValleuAlaLeu	224
Db		344	GATGGGGGGAAGGCTGTACTACCTGAAGCTGGACCTTGCTGGTAGAGTGCTGTGGCCCTGG	403
Oy		225	ArgCysIleuGluGluPheSerAlaThrAlaIaAserSerLeuGlyProGlnleuArgrLeu	244
Db		404	CGCTGCCCTGGAGAGAAATTCACCCACCTCGCGCAATTCCTCGGGCCCCAGCTCCGGCTCG	463
Oy		245	CysGlnValSerGlyLeuLeuAlaLeuArqPrgIysSerSerLeuArqIleArgrThrLeu	264
Db		464	TGCCAGGTCTGTGGGCTTGTGGCCCTGGCGGACAGGGTCTCCCTCGGATCGCACCTTC	523
Oy		265	ProTrpAlaHlsIleuLysAlaAlaProPheLeuThrTyrrPheGlyLeuPheGlnValHis	284
Db		524	CCCTGGGCCCATCTCAAGGCTGCCCCCTTCTCCACCTACTGCGACTCTCCAGGTTCAC	583
RESULT 6				
LOCUS	BFS77781			
DEFINITION	602092080P1 NCL_CGAP_CO24 Mus musculus cDNA IMAGE:4206595 5'	918 bp	mRNA	linear EST 12-DEC-2000
ACCESSION	BFS77781			
VERSION	BFS77781.1	GI:11651493		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
TITLE	1 (bases 1 to 918)			
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/.			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)  Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: Llam9767 row: P column: 20 High quality sequence start: 17 High quality sequence stop: 724.  Location/Qualifiers  1..918 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4206595" /clone_lib="NCL_CGAP_CO24" /_lab_host="DH10B (TI phage-resistant)".			
FEATURES				
SOURCE				

```

/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP library."
a 292 c 268 g 204 t 1 others

```

Alignment Scores:	
Pred. No.:	1,31e-76
Score:	913.00
Percent Similarity:	91.74%
Best Local Similarity:	87.61%
Query Match:	63.23%
Db:	12
US-09-245-198A-4 (1-284) x BF577781 (1-918)	
	Length: 918
	Matches: 191
	Conservative: 9
	Mismatches: 18
	Indels: 3
	Gaps: 0

OY		67	LeuAlaCysIleuGlyLeuIleuLeuLeuAlaValAlaSerIleuGlySerArgAlaSerIleuSer	86
Db		2	CTGGCCCTGCCTTGGCCCTCTGTGCTGTGTGTCAGACCTCGGGAGACTGGGCAACGCTGTCT	61
OY		87	AlaGlnIuProAlaGlnIuGlnIuLeuValAlaGlnIuAspGlnAspProSerGluLeu	106
Db		62	GCCCGAGGACCCTTCACGAGGAGCGCTGCACGAGGAGCCGCCGAGACCCTCCCTGAACCTG	121
OY		107	AsnProGlnInThGlnIuSerGlnAspProAlaPheProIleuAsnArgLeuValaArgPro	126
Db		122	AATCCCCAGACAGAGAANAAGCATGTGTAACCTTTCTTGGAACAACATACTAGCCGGCCT	181
OY		127	ArgArgSerAlaProLysGlyValArgLysThrArgAlaArgArgAlaIleAlaAlaHisTyr	146
Db		182	CGAAGAAGGCTCCCAAAGGCGGAAAGCGCGGCT - CGCCGAGATATTGCAGGCCATTAT	240
OY		147	GluValHisProArgProGlnIuAspGlnAspGlnIuGlnAlaGlnIuValAspGlyThrValSer	166
Db		241	GAGGTTTCATCCTCGGCCACGAGATGAGACAAAGAGGtGTGAGTAGGGACATGTAGT	300
OY		167	GlyTyrGlnIuAlaArgIleasnSerSerSerProLeuArgTyrAsnArgGlnIleGly	186
Db		301	GCGTGGGAAGAACCAAAATCAACAGCTCCACCCCTTGCGCTACGACCGCCAGATTGGG	360
OY		187	GluPheIleValaThrArgAlaGlyLeuTyrTyrLeuTyrGlyGlnValHisPheAspIu	206
Db		361	GAAATTACAGTCATCAGGAGGCTGGCTCTACACCTGTACTGTGCAGCTCACATTGATAG	420
OY		207	GlyLysAlaValaTyrLeuLysLeuAspIleuValaAspGlyValIleuAlaLeuArgCys	226
Db		421	GGAAAAGCGTGTACTACTGAGCTGAGCTTGCTGTGTAACGGTGTCGTGGCCCTGGCGTGC	480
OY		227	LeuGlnIuPheSerAlaThrAlaIaSerSerLeuGlyProGlnIuLeuArgLeuCysGln	246
Db		481	CTGGGAAGAATTCACGACCACAGCACACAGCTTCTCTGGGCCCCACACTCCGTTGTGGCAG	540
OY		247	ValSerGlyLeuLeuAlaLeuArgProGlnIuSerSerLeuArgIleArgThrLeuProTyr	266
Db		541	GTCGTGGGCTGTTCGCCCTGGCGGACAGTCT - TCCCTTGGGATTCGGCACCTCCCGCTGG	599
OY		267	AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis	284
Db		600	GCTCATCTTTAAAGCTGC - CCTTCTTAACACTTGGACTCTTTCAGATTCAC	652
RESULT 7				
LOCUS	B0707185			
DEFINITION	B0707185	948 bp	mRNA	linear EST 16-JUL-2002
ACCESSION	AGNCOURT_8353983	NIH_MGC_113	Homo sapiens cDNA clone IMAGE:6278608	
VERSION	B0707185			
KEYWORDS	B0707185.1	GI:21846084		
SOURCE	EST.			
ORGANISM	human.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

REFERENCE 1 (bases 1 to 948)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Mark Watson  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LCM2466 row: n column: 17  
 High quality sequence start: 24  
 High quality sequence stop: 550.  
 Location/Qualifiers  
 1..948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH-MGC\_113"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."  
 BASE COUNT 173 a 334 c 252 g 187 t 2 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,95e-76 Length: 948  
 Score: 909.50 Matches: 179  
 Percent Similarity: 98.35% Conservative: 0  
 Best Local Similarity: 98.35% Mismatches: 2  
 Query Match: 62.98% Indels: 2  
 DB: 14 Gaps: 1  
 US-09-245-198a-4 (1-284) x BQ707185 (1-948)  
 Oy 103 ProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArg 122  
 Db 3 CCGTCGGAATCGAATCCCGACAGAGAAAGAGCAGAGATCTCGCTTCCTGAAACGA 62  
 Oy 123 LeuValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIle 142  
 Db 63 CTAGTTCGGCTCGAAGAACTGCACCTAAAGCGCGGAAACACGCGCTCGAAGACGATC 122  
 Oy 143 AAlAlAlHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAsp 162  
 Db 123 GCAGCCCATTTATAGTTCTCCACGACCTGGACAGAGACGAGC---GCAGGtGGAC 178  
 Oy 163 GLyThrValSerGlyTyrGluGluAlaArgIleAsnSerSerSerProLeuArgTyrAsn 182  
 Db 179 GGGACAGTAGTGGCTGGGAGGAGGAGCAATCAACAGCTCCAGCCCTCGCCTACAAAC 238  
 Oy 183 ArgGlnIleGlyLupheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal 202  
 Db 239 CGCCACATCCGGAGATTATAGTCCACCGCGCTGGCTTACTACTGTCAGGTG 298  
 Oy 203 HisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeu 222  
 Db 299 CACTTTGATGAGGGGAGGCTGTCTACCTGAACCTGACTTGTGTGGATGTGTGCTG 358  
 Oy 223 AAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeu 242  
 Db 359 GCCCTGCGCTGCTGAGAGAAATCTCAGCCACATCGCGCCAGTTCCCTCGGCCCCACATC 418  
 Oy 243 ArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArg 262

Db 419 CGCCTCGCAGGCTGCTGGCGCTGGCGCCGCGCAGGCTCCCTGCGAGCCGC 478  
 Oy 263 ThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGln 282  
 Db 479 ACCCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCTACACTTTCGACTTTCAG 538  
 Oy 283 ValHis 284  
 Db 539 GTTCAC 544  
 RESULT 8  
 BO884231  
 LOCUS BO884231 940 bp mRNA linear EST 16-AUG-2002  
 DEFINITION AGENCOURT 8682031 Lupsk1.scfatic.nerve Homo sapiens cDNA clone  
 IMAGE:6197488 5', mRNA sequence.  
 ACCESSION BO884231 GI:22276239  
 VERSION BO884231.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 940)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLM13607 row: j column: 17  
 High quality sequence stop: 453.  
 Location/Qualifiers  
 1..940  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Lupsk1.scfatic.nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORE6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCGACGCGTCCG-3' and 5'-GACTAGTCTAGATCGGAGCGCGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."  
 BASE COUNT 170 a 325 c 255 g 190 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.73e-76 Length: 940  
 Score: 904.00 Matches: 183  
 Percent Similarity: 97.87% Conservative: 1  
 Best Local Similarity: 97.34% Mismatches: 0  
 Query Match: 62.60% Indels: 4  
 DB: 14 Gaps: 0  
 US-09-245-198a-4 (1-284) x BQ884231 (1-940)  
 Oy 101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeu 120

Db 1 CAGGACCCGTCGAGACTGATCCAGACAGAGAAAGCCAGATCTGCGCTTCTCTG 60  
QY 121 AAsnArgLeuValArgProArgArgSerAlaProLysGlyArgGlyThrArgAlaArgArg 140  
Db 61 AACCCAGTACGTTGGCTGCGAGAGTGCACCTAAAGCCGGAAAAACAGCGGCTCGAAGA 120  
QY 141 AAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyValGlnAlaGly 160  
Db 121 GCGATCGAGCCCATTAAGAGTTATCAGACAGCTGGACAGAGAGCGGAGCGAGAGT 180  
QY 161 ValAspGlyThrValSerGlyTyrPglGluAlaArgGlyLeuAsnSerSerProLeuArg 180  
Db 181 GTGGACGGAGACGTGAGTGGCTGGAGAGGACCAATATCAAGCTCCAGCCCTCTGGCC 240  
QY 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
Db 241 TACACCCGCGAGATCGGAGAGTTATAGTACACCCGGGCTGGGCTCTACTACTGACTGT 300  
QY 201 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220  
Db 301 CAGGTGCACCTTGTAGAGAGGAGGAGGCTGCTACCTGAAGCTGAGCTTGTGTGATGAT 360  
QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyPro 240  
Db 361 GTGCTGGCCCTGGCTGCTGGAGGAATTCAGCCACTGCGGCGCAGTTCCCTCGGCGCC 420  
QY 241 GlnLeuArgLeuCys-GlnValSerGlyLeuLeuAlaLeuArgProGlySer-SerLeu 260  
Db 421 CAGCTCCGCTCTGCGCCAGGTGTCTGGGTGTGGCCCTGCGCGCAGGGGCTCCCTGTC 480  
QY 260 TgIleArgThrLeuProThrAlaHisLeuLysAlaAla-ProPheLeuThrTyr-PheI 279  
Db 481 GGATCGGACACCTCCCTCGGGCCCATTCACAGGCTGCCCTTCTCAGCTACTTTCGG 540  
QY 279 YLeuPheGlnValHis 284  
Db 541 ACTCTTCCAGGTTTAC 556

RESULT 9  
AM763237 561 bp mRNA linear EST 04-MAY-2000  
DEFINITION ur70d09.y1 NCI-CGAP\_Mam3 Mus musculus cDNA clone IMAGE:3155633 5'  
similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS  
; mRNA sequence.  
ACCESSION AM763237  
VERSION AM763237.1 GI:7695174  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 561)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Other ESTs: ur70d09.x1  
CONTACT: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LUMN)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LUMN at:  
<http://image.llnl.gov/image/html/iresources.shtml>  
MGI:1058389  
Seq primer: -40RP from Gibco  
High quality sequence 433.  
Location/Qualifiers  
1. 561  
/organism="Mus musculus"

/strain="129, C57BL/6J, FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3155633"  
/clone\_lib="NCI-CGAP\_Mam3"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;  
Site:2: NotI; Cloned unidirectionally. Primer: oligo dr.  
Library constructed by Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 1.52e-68 Length: 561  
Score: 825.00 Matches: 161  
Percent Similarity: 90.91% Conservative: 9  
Best Local Similarity: 86.10% Mismatches: 17  
Query Match: 57.13% Indels: 0  
Gaps: 0

US-09-245-198A-4 (1-284) x AM763237 (1-561)

QY 41 SerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeuValProLeuAla 60  
Db 1 AGCCAGAGCGGAGGGGGCGCGGGGGAGCGGGACCGCCCTCTGAGCCCGCTGGCTG 60  
QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValValSerLeuGly 80  
Db 61 CTGACCTCGGGGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGluGluasp 100  
Db 121 AGCTGGGCAACGCTGTCTGCCAGAGACCTTCTCAGAGAGAGCTGACAGAGAGAGCCG 180  
QY 101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeu 120  
Db 181 CGGGAGCCCTCGAAGTGAATCCAGACAGAGAGAAAGCCAGAGTGTGACTTTCCTTG 240  
QY 121 AAsnArgLeuValArgProArgArgSerAlaProLysGlyArgGlyThrArgAlaArgArg 140  
Db 241 GAACACAGTACGTCCGCCCTCGAAGAGTGTCTCTAAAGCCGGAGCGGCTCGCCGA 300  
QY 141 AAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyValGlnAlaGly 160  
Db 301 GCTATTGACGCCCATTAAGAGTTATCAGACAGCTGGGCTCTACTACTGACTGT 360  
QY 161 ValAspGlyThrValSerGlyTyrPglGluAlaArgGlyLeuAsnSerSerProLeuArg 180  
Db 361 GTGATGGAGACGTGAGTGGCTGGAGAGGACCAAAATCAAGCTCCAGGCTCTGGAGC 420  
QY 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
Db 421 TACAGACCGCCAGATTCGGGAATNTACAGTCATCAGAGGCTGGGCTCTACTACTGACTGT 480  
QY 201 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220  
Db 481 CAGGTGCACCTTGTAGAGAGGAGGAGGCTGCTACCTGAAGCTGAGCTTGTGTAACGAT 540  
QY 221 ValLeuAlaLeuArgCysLeu 227  
Db 541 GTGCTGGCCCTGGCTGCTGCTG 561

RESULT 10  
B0671259 963 bp mRNA linear EST 15-JUN-2002  
LOCUS B0671259  
DEFINITION AGENCOURT\_8303564 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6274716  
5', mRNA sequence.  
ACCESSION B0671259  
VERSION B0671259.1 GI:21782093

KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 963)  
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution by: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM2456 row: 1 column: 13  
High quality sequence stop: 565.  
Location/Qualifiers  
1. 963  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1b="NIH-MGC\_102"  
/issue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pOT7; Site\_1: XhoI;  
Site\_2: EcoRI; CDNA made by oligo-dT priming.  
directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Library constructed  
by Ling Hong in the laboratory of Gerald M. Rubin  
(University of California, Berkeley) using ZAP-CDNA  
synthesis kit (Stratagene) and Superscript II RT (Life  
Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 194 a 326 c 260 g 183 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,12e-66 Length: 963  
Score: 809.00 Matches: 160  
Percent Similarity: 99.388 Conservative: 0  
Best Local Similarity: 99.388 Mismatches: 1  
Query Match: 56.028 Indels: 1  
DB: 14 Gaps: 0

US-09-245-198a-4 (1-284) x B0671259 (1-963)

OY 124 ValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAla 143  
|||||  
Db 2 GTTCGGCTCGCAGAGTCACCTAAAGCGCGAAAAA-CGGCTCGAAGCGATCGCA 60  
|||||  
OY 144 AlAHistyrLysValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163  
|||||  
Db 61 GCCCATTTATGAATCTCATCCACCTGGACAGACGACGACGAGGAGGTGTGACGGG 120  
|||||  
OY 164 ThrValSerGlyTrpGluGlnAlaArgIleAsnSerSerProLeuArgTyrAsnArg 183  
|||||  
Db 121 ACAGTAGTGCGTGGGAGGAGCAAGCATCAACAGCTCCAGCCCTCGCCCTCAACACCGC 180  
|||||  
OY 184 GluIleGlyLysPheIleValIleThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHis 203  
|||||  
Db 181 CACATGGGGAGTTTATAGTCACCCGGCGGTGCTACTACTGTTACTGTCAAGTGCAC 240  
|||||  
OY 204 PheAspGluGlyLysAlaValIleTyrLeuLysLeuAspLeuValAspGlyValLeuAla 223  
|||||  
Db 241 TTGTAGTAGGGAGGAGGTGTCTACTGAAGCTGACTGTGTGTGATGTGTGTGCGCC 300  
|||||  
OY 224 LeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg 243  
|||||  
Db 301 CTGCGGTGCTGGAGGATTTCTACGCCACATGCGGCCAGTTCCCTCGGGCCCAAGCTCCGC 360  
|||||

OY 244 LeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThr 263  
|||||  
Db 361 CTCTGCCAGGTGTCTGGGCTGTGGCCCGGCGGAGGAGGCTCTCCCTGGAGATCCGACCC 420  
|||||  
OY 264 LeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283  
|||||  
Db 421 CTCCCTGGGCCCATCTCAAGGCTGCCCTTCCTCACTACTGTGCGACTCTTCAGGTT 480  
|||||  
OY 284 His 284  
|||  
Db 481 CAC 483

RESULT 11  
BM521213 1071 bp mRNA linear EST 12-MAR-2002  
LOCUS  
DEFINITION AGENCOURT\_6633046 NIH-MGC\_115 Homo sapiens CDNA clone IMAGE:5752561  
5', mRNA sequence.  
ACCESSION BM521213  
VERSION BM521213  
KEYWORDS BM521213.1 GI:19371592  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 1071)  
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution by: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LAM12786 row: P column: 02  
High quality sequence stop: 656.  
Location/Qualifiers  
1. 1071  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1b="IMAGE:5752561"  
/clone\_1b="NIH-MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27; and 1 male testis, age 69. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH-MGC Library."

BASE COUNT 226 a 346 c 279 g 220 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.02e-62 Length: 1071  
Score: 765.00 Matches: 147  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 52.98% Indels: 0  
DB: 14 Gaps: 0

US-09-245-198a-4 (1-284) x BM521213 (1-1071)

OY 138 AlaArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAla 157  
|||||  
Db 12 GCTCGAAGAGCATCCGACGCCCATTTATGAAGTTCAATCCAGACACTGGACAGGAGCGG 71  
|||||

Oy	158	G1A1AG1ValAlAspGlyThrValSerGlyTrpArgLysIleAlaArgLysSerSerSer	177
Db	72	CAGCAGAGGTGGACGGGACAGTAGTGGCTGGGAGAAGCACAGATCAACAGCTCCAGC	131
Oy	178	ProlenAgtYrAsnArGlnIleGlYgLuPheIleValIThrArgAlaGlyLeuTyrr	197
Db	132	CCTCGGCCCTCAACACCAGCATCGGGAGATTATAGTCACCCGGGCTGGGCTCTACTAC	191
Oy	198	LeuTYrCySGlnValHisPheAspGluGlyLysAlaValTYrLeuLYSLeAspLeuLeu	217
Db	192	CTGTACTCTCAGAGTGTCACCTTTGATGAGGGGAAGGCTGCTCACTGAACTGGACTTGC	251
Oy	218	ValAspGlyValIleAlaLeuAlaLeuArgGlySerLeuGluGluPheSerAlaThrAlaLysSer	237
Db	252	TGGATGGGTGGCTGGGCCCTCGGCTGGCTGGAGGAATTCACGCCACGCGGAGTTCC	311
Oy	238	LeuGlyProGlnLeuAlaArgLeucYSglnValSerGlyLeuLeuAlaLeuArgProGlySer	257
Db	312	CTCGGGCCCCAGACTCGGCTCTGCCAGGATGTCGTGGGCTGTGGCCCTCGGCCAGAGTCC	371
Oy	258	SerLeuArgIleArgTYrThLeuProTrpAlaHisLeuLYSalaalaaProPheLeuThTYr	277
Db	372	TCCCGCCGAGTCCGACCCCTCCCGGCGCCATCTCAGAGCTGCCCCCTTCTCACCCTAC	431
Oy	278	PheGlyLeuPheGlnValHis 284 	
Db	432	TTCCGACCTCTCCAGGTTCCAC 452 	
RESULT 12			
LOCUS	B0674188	951 bp	mRNA linear EST 15-JUL-2002
DEFINITION	AGENCOURT_8354100 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275664		
ACCESSION	B0674188		
VERSION	B0674188.1 GI:21785022		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 951) NIH-MGC http://mgc.ncl.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: http://image.lnl.gov plate: LHCW2459 row: d column: 01 High quality sequence stop: 335. Location/Qualifiers 1..951 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6275664" /clone_id="NIH_MGC_102" /tissue_type="epidermoid carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAGC(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	165 a	286 c	315 g 184 t others
ORIGIN			

Alignment Scores:	2,356-61	Length:	951
Pred. No.:	753.00	Matches:	152
Score:	95.62%	Conservative:	1
Percent Similarity:	95.00%	Mismatches:	5
Best Local Similarity:	52.15%	Indels:	2
Query Match:	14	Gaps:	0
DB:			

  

US-09-245-198A-4 (1-284) x BQ674188 (1-951)	
QY 124	valaA-gProaA-gA-gSerA-lA-pRoLysG-lY-aRgLyS-thrA-gAlA-aA-gA-rAlA-lleA 143
Db 2	gTTCGGCCCTCCGAGAAAGTGCACACTAAAGCCGGAAACACAGGGCTCGAAAGAGCATGCCA 61
QY 144	AlAh1StYrG1uVAlH1S1PrOaRgPrOblYg1nAspRyAlaG1nAlaG1VAlAspRy 163
Db 62	GCCCATTTATGAAGTTCATCCACGACTCTGACAGCAGGAGCGAGCGAGGTGTGACGG 121
QY 164	ThrValSerG1YTPRg1uG1uAlAaRg1leAsnSerSerSerProLeuAArgTyAsnArg 183
Db 122	ACAGGAGATGGCTGGGAGAGAACCGAGATCAACAGCTCCACCCCTCTCGCTACACCGC 181
QY 184	G1u1leG1uPhe1leVal1ThrArg1aG1LeuTyTYrLeuTyCYsG1nVAlH1S 203
Db 182	CAGATCGGGGAGATTATAGTCAACCCGGGCTGGGCTCTACTACGTACTGACAGTGCAC 241
QY 204	PheaSpR1uG1uLySAlAValTYrLeuLysLeuAspLeuValAspG1VAlleuA 223
Db 242	TTTGTATGAGGGGAAAGGCTGTCTACCTGAAGCTGCGCTGTGTGAGTGTGTGGCC 301
QY 224	LeuAArgCYsLeuG1uG1uPheSerA1aThrAlAa1aSerSerLeuG1uProG1nleuArg 243
Db 302	CTGGCCTGCTGTGAGGAATTTCTCACCCACTGGCGGCAATTCCTCTGGGGCCCACTCCGC 361
QY 244	LeuCYsG1nValSerG1yLeuLeuAlAa1eAaArg-ProG1ySerSerLeuAArg1leAArgTh 263
Db 362	CTCTCGACAGAGTGTCTGGGCTGTGGCCCTCGGGGACAGGTCCTCCCTGGGGATCCCCAC 421
QY 263	r1eUpRoTTPAlAh1S1eLysAlAAlA1-ProPheLeuThTYrPheG1yLeuPhe 281
Db 422	CCTCCCGCGGGGCCCATTTAAAGGGTGGCCCTTTCTTCACCTATTGTGGAAATTTT 477

  

RESULT 13	785 bp	mRNA	linear	EST 25-SEP-2001
B1762908				
LOCUS	603047966F1	NIH_MGC_116	Homo sapiens cDNA clone IMAGE:5188168	5',
DEFINITION		mRNA sequence.		
ACCESSION	B1762908			
VERSION	B1762908.1	GI:15754486		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 785)			
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
COMMENT	Unpublished (1999)			
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Contact: Robert Strausberg, Ph.D.			
	Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>			
	Tissue Procurement: Life Technologies, Inc.			
	cDNA Library Preparation: Life Technologies, Inc.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)			
	DNA sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LNL at:			
	<a href="http://image.lnl.gov">http://image.lnl.gov</a>			
	Plate: L1AM1470 row: C column: 17			
	High quality sequence stop: 647.			
	Location/Qualifiers			
FEATURES	1..785			
source	/organism="Homo sapiens"			





Db	137	GGGAGCCGGGACCGGCCCTCTGTATCCCGCTCGGCGTGGACCTGGGCTTGCTTCG	
Oy	68	aCyLeuclyleutleuAlaValSerLeuclyserTrgAlaSerLeuSerAlaGl	88
Db	197	CTGCCTGGCCTCTCTGCTGGCGCGTGGTCATGTTGGGGACCGGGCATCTGTCGGCCA	256
Oy	88	nGIAPrAlaInglIngluleuValAlaInglInAspInAspInAspProSerGIUleuAnPr	108
Db	257	GGACCTGGCCCCAGAGAGAGCTGGTGGCGAGAGAGACCAGCACCGCTCGGAATGAATGC	316
Oy	108	oGIInThGlUGlUSerGINaspproAlaPhePheluAnsArGLueValAArgproAr	128
Db	317	CCAAcAGAAAGAACCAAGATCTGGCCCTTCTCTGAACCGACTAGTTCGGCCCTCGAG	376
Oy	128	gSeFAlProLySGIARqLysTrnArgAlaArgAlaIleAlaIahISTYrGLuVA	148
Db	377	AAGTGCACCTTAAGCGCCGGAACACAGGGCTCGAAGACGATTCGAGCCCATTAAGAAGT	436
Oy	148	IHisProArGrProGLyGlnAspgLYalAGlnAlaGLyValAspGLy	163
Db	437	TCAACCAcGACCTGAGACGAGCAGCGCCAGCA-----GATGGA	476
<b>RESULT 15</b>			
LOCUS	AKO20909	1033 bp	mRNA linear HTC 19-JAN-2002
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930030D13:tumor necrosis factor (ligand) superfamily, member 12, full insert sequence.		
ACCESSION	AKO20909		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (Strain:C57BL/6J) adult retina cDNA to mRNA, clone:A930030D13. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
<b>ORGANISM</b>			
REFERENCE	1	Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)	
JOURNAL MEDLINE PUBMED	99279253		
REFERENCE	2	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	
AUTHORS			
JOURNAL MEDLINE PUBMED	20499374		
REFERENCE	3	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komno,H., Akiyama,D., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M., Sumi-N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamanoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiyagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer Genome Res. 10 (11), 1757-1771 (2000)	
TITLE			
JOURNAL MEDLINE PUBMED	20530913		
REFERENCE	4	Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Aikawa,T., Hara,A., Fukunishi,Y., Komno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Komodo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Platschmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Knehl,P., Lewis,S., Matsuno,Y., Nikaido,I., Pesole,G.,	
AUTHORS			

[illegible]



RS 5/15/03

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:59 ; Search time 54.1218 Seconds  
(without alignments)  
1081.216 Million cell updates/sec

Title: US-09-245-198A-4  
Perfect score: 1444  
Sequence: 1 MSLLDFRISARLLPLPRSLG.....PMAHLKAPFLTYGFLQVH 284

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	8.0	325	5	Q9V5G2
2	103	7.1	565	16	Q9K766
3	101	7.0	643	16	Q9K217
4	100.5	7.0	375	16	Q9RRH5
5	100	6.9	850	11	Q9J715
6	99.5	6.9	378	16	Q9AAB9
7	99	6.9	614	4	Q8WU14
8	99	6.9	835	4	Q9NKR9
9	99	6.9	878	4	Q9UFU7
10	98	6.8	206	16	Q9S2M5
11	97.5	6.8	1560	4	Q9GJP2
12	97	6.7	408	10	Q8S515
13	97	6.7	926	4	Q9N7A0
14	97	6.7	1058	4	Q9Y4G2
15	96	6.6	629	10	Q8S1A6
16	95.5	6.6	536	4	Q9HB96

17	95.5	6.6	655	16	Q9FBR7	Q9fbr7 streptomyc
18	94.5	6.5	748	5	Q8T2Y0	Q8t2y0 trypanosoma
19	94.5	6.5	1696	11	Q9WTR8	Q9wtr8 ratius norv
20	94	6.5	937	16	Q9JUD1	Q9jud1 streptomyc
21	93.5	6.5	776	16	Q9RJ01	Q9rj01 streptomyc
22	93.5	6.5	854	16	Q9FZP0	Q9fzp0 streptomyc
23	93.5	6.5	1038	10	Q9AS09	Q9as09 oryza sativ
24	93	6.4	507	2	Q9RM75	Q9rm75 streptomyc
25	93	6.4	1043	10	Q93VC8	Q93vc8 oryza sativ
26	92.5	6.4	484	11	Q9CTA7	Q9cta7 mus musculu
27	92.5	6.4	614	4	Q9NW41	Q9nw41 homo sapien
28	92.5	6.4	707	11	Q925J8	Q925j8 mus musculu
29	92.5	6.4	1511	13	Q91BE2	Q91be2 fugu rubrip
30	92	6.4	351	12	Q85302	Q85302 orf virus.
31	92	6.4	566	16	Q70005	Q70005 streptomyc
32	92	6.4	845	4	Q96H68	Q96h68 homo sapien
33	92	6.4	977	4	Q8TE71	Q8tel1 homo sapien
34	91.5	6.3	212	16	Q86312	Q86312 mycobacteri
35	91.5	6.3	221	10	Q8RYZ6	Q8ryz6 oryza sativ
36	91.5	6.3	664	16	Q9HYT9	Q9hyt9 pseudomonas
37	91.5	6.3	3169	5	Q917V8	Q917v8 drosophila
38	91	6.3	681	16	Q9HUZ6	Q9huz6 pseudomonas
39	91	6.3	1217	4	Q60336	Q60336 homo sapien
40	90.5	6.3	449	2	Q93HD5	Q93hd5 streptomyc
41	90.5	6.3	760	11	Q8VD14	Q8vd14 mus musculu
42	90.5	6.3	1100	2	Q937L3	Q937l3 delinococcus
43	90.5	6.3	1329	16	Q9CD30	Q9cd30 mycobacteri
44	90	6.2	285	2	Q33629	Q33629 saccharopol
45	90	6.2	422	16	Q91OB9	Q91ob9 streptomyc

## ALIGNMENTS

## RESULT 1

ID	Q9V5G2	PRELIMINARY;	PRT;	325 AA.
AC	Q9V5G2:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	CG12919 protein.			
GN	CG12919.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,			
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,			
RA	Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matedel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relneert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Splier E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskis R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003831; AAF58848.1; -;  
 DR Flybase: FBgn003483; CG12919.  
 DR InterPro: IPR000478; TNF\_family.  
 DR SMART: SM00207; TNF\_1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PSS0049; TNF\_2; 1.  
 SO SEQUENCE 325 AA; 36862 MW; 6E5CCB6964F1A3A CRC64;

Query Match 8.0%; Score 116; DB 5; Length 325;  
 Best Local Similarity 25.4%; Pred. No. 0.07;  
 Matches 57; Conservative 32; Mismatches 97; Indels 38; Gaps 10;

QY 82 RASLSAQEPAQELVAE---EPDPSSELNPTESQDPAFLNLR-----LYRPR 127  
 DB 118 RRSRSIADVRNEQNTQGHTELDQKSSNEATSKES--PAPLHRRRMSRHRLVLRKG 175  
 QY 128 RSAPKRRKTRARAIANHEVHPRPQDGAQAGVDTSVSGMEARLNSSPLRYNQIGE 187  
 DB 176 ESLLSARSDSR--AAHPLLSRRHQSQM--GYHDMYIGNNENNSYOG--HFQRRDGV 231  
 QY 188 FIVTRAGLYLYCOV-----HEDEKAVYLYKLDLVGVALRCLCEFSATASSLGPOL 242  
 DB 232 LFTVTNGLYVVAQICYNNSHDQNGFIVF-----QGDPTEFLOCLN---TVPNMPHKV 281  
 QY 243 RLCOVSGLLALRPSSSLRIRTL--PWAHLKAPFLTYRGLPOV 283  
 DB 282 HTCHTSGLIHLERNERIHLDIHDRNAVALREGNNSYFEIRKV 325

## RESULT 2

Q9KY66 PRELIMINARY: PRT; 565 AA.  
 AC Q9KY66: 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Putative peptidase.  
 GN SC04798 OR SC063A.09C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID-1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Brown S.P., Harris D.;  
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE-97000351; PubMed-8843436;

RA Redenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb streptomycetes coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Thomson N.R., James K.D., Harris D., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL356832; CAB92661.1; -;  
 DR InterPro: IPR002886; Peptidase\_M37.  
 DR Pfam: PF01551; Peptidase\_M37; 1.  
 SO SEQUENCE 565 AA; 58070 MW; 7D0418D480C6A284 CRC64;

Query Match 7.1%; Score 103; DB 16; Length 565;  
 Best Local Similarity 21.2%; Pred. No. 1.7;  
 Matches 70; Conservative 33; Mismatches 99; Indels 128; Gaps 16;

QY 2 SLIDFELSA---RLPLRSLGSDGAVRQAPPPMARRGRRGRGEGTALL-- 56  
 DB 290 ALLDDEEATPDARIPVARR--GSRAARSRRKP-----AKRS-----ALLTI 331  
 QY 57 -VPLALGL--GLALACILLLAVVSLGSRASLSAQEPAQEE----- 94  
 DB 332 AVPSACVMSVAGIAAASVGLTG--DEGTETASAPDPCGAELAPVPSAANKKLDTOIL 389  
 QY 95 -LVAEED-----QDPSELNPTESQDPAFLNLRVPRNSAPYGRKTRARRAIANY 146  
 DB 390 SLAAGADDFADRASRTOERIDLKAEODAEKRAAQEAARKERLRPRFALPVKQHGSLAVY 449  
 QY 147 EVNRPQGDG-----AQAGVDTSVSGMEARLNSSPLRYNQIGE 187  
 DB 450 -----GQAGINMSSHTGIDFPVLOGTIVMAKTDGIVR-----TQFNSAYGN 491  
 QY 188 FIVTRA--GLYLYCOVH-----FDEGKAVYLYKLDLVGVALRCLCEFSATAAS 236  
 DB 492 MMIVTAKDGTETWYCHLSQVPSGTYVAKGDAI-----AYSGDSGV 533  
 QY 237 SLGPOLRLCOVSGLLALRPSSSLRIRTLPW 266  
 DB 534 STGPHLH-----FEVPRAGSSIDPLPW 556

## RESULT 3

Q9KZ17 PRELIMINARY: PRT; 643 AA.  
 AC Q9KZ17: 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Hypothetical protein SC02220.  
 GN SC02220 OR SC10B7.15.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID-1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,



```
Db 601 LGSPTTPPALNGNPPRSPPEEPERRRAGPTRENGSSGTPKLIORALRCGTA-----LLA 654
QY 76 VVSLG-----SR---ASLSAQEPAO-----BELVAEEDQDPELNPQTEESODPA 117
Db 655 SLGSGRDLQPPGSLSRGSPPTAPPQMPSPCPPELPTPLIRLSQTTPTDAHSSPTPG 714
QY 118 PFLNRL-----VPRPRSAFKGKTRARRAIAHYEVHPRPGDGAQAQVGDCT 164
Db 715 PLLLDLGVSPGSPSAKSPRETRGR-----TVSPPEISNSAPETPCT 758

RESULT 6
Q9AAB9 PRELIMINARY; PRT; 378 AA.
ID 09AAB9;
AC 09AAB9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hlyd family secretion protein.
GN CC0683.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter;
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; Pubmed=11259647;
RA Nielsen W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Elsen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Polocka J., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utezhack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005743; AAK22668.1; -
DR TIGR; CC0683; -
DR InterPro: IPR003997; RTXDP.
DR PRINTS; PR01490; RTXTOXIND.
KW Complete proteome.
SQ SEQUENCE 378 AA; 39659 MW; 88B818C437C15CBC CRC64;

Query Match 6.9%; Score 99.5; DB 16; Length 378;
Best Local Similarity 25.1%; Pred. No. 2.1;
Matches 73; Conservative 29; Mismatches 106; Indels 83; Gaps 13;

QY 19 LGSRDGAVRQAQPPAPMAARRSQRRGR--GEGTALLVPLALGLALACLGILLA 75
Db 36 LAAROGGIVANVH-----REGDRVVGQPIATLTLSALRGDSFVLSR--- 81
QY 76 VVSLGSAASLSAQEP--QBELVAEEDQDPELNPQTEESODPAPELNPVPRRS--AP 131
Db 82 --SLGAQSSAAGSRAAATQAALAAESRQ-----LVQRRALNRELAEAKRSALQ 139
QY 132 KGRTRR-----RAIAHYEVHPRPGDGAQAQVGDCTVSGMEARINSSPLRYRQI 185
Db 130 TERITLARAFAVARETTIAAGGFLSPR-----LQARRSAELAVQDEASTLSQALSTYRQI 185
QY 186 GEFIVTRAGLYLYLCOVHDEGRAVYLKLDL---LVYDGVIALRCLREFSATASSLGPOL 242
Db 186 GEV-----DARLAIPIDQLQAARAAASTLAGLEG-QATQVYEAQGRYV 227
QY 243 RLCOVSGLLA-----LRPGSLRIRITPMALHKAAPL 275
Db 228 VVATVAGRVAAALPVEAQIVGVGAANAVALTPGDSALVAEL-VAPSRAGGV 277

RESULT 7
Q8WU14 PRELIMINARY; PRT; 614 AA.
ID 08WU14
```

```
AC Q8WU14; Q9BR73;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Histone deacetylase 7A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RA Strausberg R.;
RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE OF 339-614 FROM N.A.
RC TISSUE-B-CELL;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020505; AAH20505.1; -
DR EMBL; BC006453; AAH06453.1; -
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 614 AA; 66186 MW; 2B638D8C666B502 CRC64;

Query Match 6.9%; Score 99; DB 4; Length 614;
Best Local Similarity 24.2%; Pred. No. 4.2;
Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

QY 11 RLPLPRSLGSRGGAVR-----QAQPPAPMAAR-----RSQRRGR- 47
Db 63 RQIPSAEDLTDDGGGQGVDDGLREHRELGHGPEARGAPALQOHPOVYLMEDQRLAGRL 122
QY 48 -RPPGTALLVPLALGLALACLGILLAVVSLGSRASLSAQEPQBELVAEEDQDPEL 106
Db 123 PRSTGTVLLPLAAGGHPLS-----RAQSPAAPASLSAPPASQARVLSSETPART 177
QY 107 NPQT-----EESODP-----APFLNRYVPRRSAPKGRTRAR 139
Db 178 LPFTTGILYDSVLMKHCQSCGDSRHPEHNGRIQSIWSRLQERGLMSQCCLRGRAASLE 237
QY 140 RAIAHYEVH 149
Db 238 ELQSVHSERH 247

RESULT 8
Q9NKR9 PRELIMINARY; PRT; 855 AA.
ID 09NKR9;
AC 09NKR9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Histone deacetylase 7.
GN HDAC7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-CERVICAL CARCINOMA;
RA Li S., Fischle W., Verdin E., Walsh M.J.;
RT "A novel class II HDAC is associated with the transcriptional
homeodomain repressor CCAAT displacement protein.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF239243; AAF63491.1; -
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 855 AA; 92363 MW; 3157426AE1913DB6 CRC64;
```







```

Oy 117 -----APFLNR--LVPRRSAPKGRKTRARARATAAHVEVHPRPGDGAQGVDPGVSG 167
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 337 YRGTGSPRGRLHRHRSQEPKRCISGLALDQ-----CVPSPGRQAQAPPS---QG 384
Oy 168 WEARL-----NSSSPLRYNRQIGEFIVTRAGLY-----YLCOVHFDEGKAVYLKLDL 217
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 385 HKSFRVYHRQMGISNPFRLMKLG--TVERRGAMGIMKELFCELSPLEFR-LYLSNE-- 439
Oy 218 VDGVALRCLEFESATASSIGP-----QLRICOVSGLLALRPGS 257
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 440 -----EHTCVENCSSLRCESVGPAPHSDDRFEFLVFSGKLLALRASS 479

RESULT 14
Oy 09Y4G2 PRELIMINARY; PRT; 1058 AA.
ID 09Y4G2;
AC 09Y4G2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE KIA0356 protein (Fragment).
GN KIA0356.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.
RL DNA Res. 4:141-150(1997).
DR EMBL; AB002354; BAA20813.2;
DR InterPro; IPR002219; DAG_PE_bind.
DR InterPro; IPR000532; Glucagon.
DR InterPro; IPR001849; PH.
DR InterPro; IPR004012; Run.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PR00275; GLUCAGON.
DR SMART; SM00109; C1; 1.
DR SMART; SM00233; PH; 2.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
FT NON_TER 1 1
SQ SEQUENCE 1058 AA; 117671 MW; CD1AD1869C8C9D1 CRC64;

Query Match 6.7%; Score 97; DB 4; Length 1058;
Best Local Similarity 24.2%; Pred. No. 12;
Matches 69; Conservative 35; Mismatches 125; Indels 56; Gaps 13;

```

```

Db 572 -----EHTCVENCSSLRCESVGPAPHSDDRFEFLVFSGKLLALRASS 611
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|

RESULT 15
Oy 08S1A6 PRELIMINARY; PRT; 629 AA.
ID 08S1A6;
AC 08S1A6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE B114B07.16 protein.
GN B114B07.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(Ga3) genomic DNA, chromosome 1, BAC
RT clone:B114B07."
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP003334; BAB90006.1;
SQ SEQUENCE 629 AA; 68639 MW; 724C08A73DCB785D CRC64;

Query Match 6.6%; Score 96; DB 10; Length 629;
Best Local Similarity 20.8%; Pred. No. 7.7;
Matches 67; Conservative 33; Mismatches 116; Indels 106; Gaps 11;

```

```

Oy 12 RPLPRSLGSRGAVRQAQPPAPMAARSORRRGRGEPG-----AL 55
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 107 RRPYRPPREAHVLRQAQPPRRRVRGRRRRRRAVPDVRDGDGSRSLRRRRRLS 166
Oy 56 LVPALGLG-----LALACL----- 70
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 167 LADTAVAGLSVDVARRRHGCGRSGGGGGGGGGLGGASHATPALTALIPRGTAPCRG 226
Oy 71 -GLLAVVSLGSRAS-----LSAQEPQAEELVAEPDQDSELPQTEESODPAPFLNRLVR 125
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 227 RGAALAPGSLGRRGCGORLACVGAERADDEGRDPGVEDRGRDGEPAEEVALVY- 285
Oy 126 PRRSAPKGRKTRARARATAAHVEVHPRPGDGAQGVDPGVSGWEARINSSSPLRYNRQI 185
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 286 ---ERGGGGRVYRRRRHHWYRERERERESAAGD-----RRRT 324
Oy 186 GEIVTRAGLYLYCOVHFDEGKAVYLKLDLVLDVLAIRCLEFESATPAASSLGPOLRLC 245
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 325 LEYVAT--CVHGLVCPHIAATFRSL-----AACVLPYVAVMNARALRSIGPG-RQC 372
Oy 246 OVSGLALRPGSSLRIRLPMA 267
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
Db 373 -VAG-----RGSKSPVPTRDWS 388

Search completed: March 31, 2003, 07:46:23
Job time : 62.1218 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 07:44:21 ; Search time 28.4558 seconds  
(without alignments)  
959.459 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARRLLPLPSRLG.....PMAHLKAAPLTYFGRLQVH 284

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*

1: pir1:\*\n2: pir2:\*\n3: pir3:\*\n4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106.5	7.4	254	2	4-1BB ligand - hum
2	104	7.2	1323	2	N-methyl-D-asparta
3	100.5	7.0	375	2	hypothetical prote
4	99.5	6.9	378	2	HLyd family secret
5	99	6.9	878	2	hypothetical prote
6	98	6.8	206	2	probable membrane
7	97	6.7	441	2	mitosis-specific c
8	93.5	6.5	776	2	probable cation-tr
9	93.5	6.5	1657	2	hypothetical prote
10	93	6.4	310	2	hypothetical prote
11	92	6.4	351	2	ORF5 protein - Orf
12	92	6.4	566	2	probable two-compo
13	91.5	6.3	212	2	hypothetical prote
14	91.5	6.3	664	2	hypothetical prote
15	91	6.3	681	2	hypothetical prote
16	91	6.3	1217	2	2,4-dienoyl-CoA re
17	90.5	6.3	814	2	hypothetical prote
18	90.5	6.3	1329	2	cadherin-15 precu
19	90	6.2	660	2	conserved hypothet
20	89.5	6.2	886	2	probable inhibitor
21	88.5	6.1	210	2	hypothetical prote
22	88.5	6.1	278	2	hypothetical prote
23	88.5	6.1	281	2	fas ligand - rat
24	88.5	6.1	998	2	probable two-compo
25	88	6.1	531	2	conserved hypothet
26	87.5	6.0	492	2	hypothetical prote
27	87	6.0	197	1	tumor necrosis fac
28	87	6.0	204	1	lymphotoxin - bovi
29	87	6.0	445	2	hypothetical prote

30	87	6.0	2124	2	H83357	probable non-ribos
31	86.5	6.0	399	2	F83633	hypothetical prote
32	86.5	6.0	439	2	B70629	hypothetical prote
33	86.5	6.0	755	2	B75346	probable competent
34	86.5	6.0	762	2	E98121	hypothetical prote
35	86.5	6.0	764	1	S14113	1-phosphatidylinos
36	86.5	6.0	810	2	D95256	Arp-dependent Clp
37	86.5	6.0	887	2	AC0535	ClpB-like protein
38	86	6.0	139	2	B69953	hypothetical prote
39	86	6.0	565	2	G98331	probable oligopept
40	86	6.0	568	2	E83325	probable chemotaxi
41	86	6.0	777	2	A87309	hypothetical prote
42	86	6.0	839	2	F75518	hypothetical prote
43	86	6.0	933	1	B48349	glycoprotein B pre
44	85.5	5.9	563	2	C98227	hypothetical prote
45	85.5	5.9	563	2	AB3059	hypothetical prote

## ALIGNMENTS

RESULT 1  
138427  
4-1BB ligand - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 21-Jul-2000  
C:Accession: 138427  
R:Alderson, M.R.; Smith, C.A.; Tough, T.W.; Davis-Smith, T.; Armitage, R.J.; Falk, B.  
Eur. J. Immunol. 24, 2219-2227, 1994  
A:Title: Molecular and biological characterization of human 4-1BB and its ligand.  
A:Reference number: 138426; MUID:94374434; PMID:8088337  
A:Accession: 138427  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-254 <RES>  
A:Cross-references: EMBL:003398; NID:9571322; PIDN:AAAS3134.1; PID:9571323

Query Match  
Best Local Similarity 26.5%; Pred. No. 0.32;  
Matches 74; Conservative 31; Mismatches 93; Indels 81; Gaps 13;

QY	32	PPAPMAARRSGRRGCEPTALLVPLALGGLLAVVSL-GSRASL-SAGE	89
DB	16	PPAP-----RARCVRPL-WALVAGLLLLLAACAVFLACPMAVGARSAFGSAAS	67
QY	90	PAQELVAEEDDPSELINPQTEESQDPAPFLNRLVRRRSAPKGRTRARRAIAHYEVH	149
DB	68	PLIRE-----GPGLSP-----DDPAGLLDL-----RQGMFAQLVAQNVLL-	102
QY	150	PRPGDGAQAGVGVSGMEE--ARINSSPLRYNRQIGEFYTRAGLYTYLCOVHFE	206
DB	103	-----IDGPLSWISDPGLAGVSLTGGLSTKEDTKELVAKAGVYVFFQ-----	146
QY	207	GRAVYLKLDLVDG-----VLALRCLEEFSAATASSIGPOLRLCOVSG-----	249
DB	147	-----LELRVVAGSGSSVSLALHQLPLRSAAGAAALATVLDLPASSEARNSAFGFG	201
QY	250	-LLALRPSSSLRIKTLPMHLKAAPL-----TYGELFOV	283
DB	202	RLHLISAGQRLGVLHTEARARHAWLTOGATVGLFRV	240

RESULT 2  
S27224  
N-methyl-D-aspartate receptor epsilon-4 chain - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 25-Feb-1994 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
C:Accession: S27224  
R:Iikeda, K.; Nagasawa, M.; Mori, H.; Araki, K.; Sakimura, K.; Watanabe, M.; Inoue, Y.  
FEBS Lett. 313, 34-38, 1992  
A:Title: Cloning and expression of the epsilon-4 subunit of the NMDA receptor channel  
A:Reference number: S27224; MUID:93050214; PMID:1385220  
A:Accession: S27224







	Best Local Similarity	24.5%	Pred. No. 11:	
	Matches	66;	Conservative	24; Mismatches 99; Indels 80; Gaps 14;
OY	PPAPMAARRSOR----	RRCRRGE-----	GTAALLYPLALGLIACILAVVSLGS--	81
Db	2 PPAPTVMAMSTPPEVRRLRGLPRFVSQVLMLQLAIAAGAAVATGTLFLA--	PLGDQL		59
OY	82 -----RASLSAQEPAGBELVAEEDQDPSSELNPQEESSODPAPIFNLRVPRRSAPKRRK			135
			:	
Db	60 DDQAMRRRLAIATQTAAQQPVAVRD-----	LRTTRTPANCPVQRE		98
OY	136 T-RARRATAAHVEV-----	HPRGODGAQAGVD--	GVSGWEARINSSS-----	177
Db	99 AERVREARAEIVYMNDQGVMSHTDPERIGEIVYSTPGCALAGERMEIDDLGGRSA			158
OY	178 ----PLRYNROIIEFI-VTRAGLYLYLCQ----	VHEDECKAVY----	IKLDELVDGVIALR	225
Db	159 RKGVPALRGDD--GETVGAVSVGIAYDSVARRLIHAPGLFAVAGALAVGALAISWIIISR			216
OY	226 CLEESATFAASSLGQRLCOVSGILLALR			254
	:	:	:	:
Db	217 VOROTRIAFS-----	DIAGLIAER		236

```

RESULT 13
A70611
hypothetical protein Rv1219c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70611
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fieldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, R.; Skellton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70611
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-212 <COL>
A:Cross-references: GB:593777; GB:AL123456; NID:93261726; PIDN:CAB07841.1; PID:e1299832.
C:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv1219c

Query March 6.3%, Score 91.5; DB 2; Length 212;
Best Local Similarity 24.5%; Pred. No. 4.1;
Matches 53; Conservative 26; Mismatches 78; Indels 59; Gaps 9;

QY 46 GRGEPGTALLVPLAIGL---ALACIGLLAVNSIGSRASISAOEPAOELVAEE-- 99
||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-|||
Db 21 GRHG-----FGVGLRAIAEAGVSAAALYVHFGSGEGL---KKACDDFFAEIIR 66

QY 100 -----DQPSLEINPQTESQDPAEFLNRLVPRRSAPKGRKTRARRAIAHYEHR 151
||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-|||
Db 67 SSKAALKSNPPTTLMQMAEIESYAPLMAYLVSMQSGELAKMLQKMI----- 117

QY 152 PGQAGAGCGVGYSGWEARINSSPLRYNRQIGEFI-VTRAGIYLYXCVHPD----- 205
||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-|||
Db 118 ---DNAEYLD-----EGVRAGTYVPSRDPRARRAFLAITGGGGFLLYLDIMHENPTDLR 168

QY 206 -----EGKAVYLKLDLIDVGLALRCL-EEFSATA 234
||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-||| 1-|||
Db 169 AALRDYAHMDVLPISLEVYTEGLADRAMYEAFLEA 204

RESULT 14
D83231
hypothetical protein PA3305 (imported) - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83231

```

R:Stover, C.K., Pham, X.Q., Errington, A.L., Mizoguchi, S.D., Warrenner, P., Hickey, M.J., Adaman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lory, S., Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: D83231  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <STO>  
A:Cross-references: GB:AE004753; GB:AE004091; MID:g9949433; PIDN:AA606693.1; GSPDB:GM1000000000  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3305

```

Query March 6.3% Score 91.5; DB 2; Length 664;
Best Local Similarity 25.8%; Pred. No. 15;
Matches 77; Conservative 28; Mismatches 117; Indels 77; Gaps 12;

Oy 28 ROAOPAPMAAR-----SOR-----RRRRGEGTALLVPLAAGLALACGLLL 74
      ||| : ||| ||| ||| |||
Db 194 ROPOGGLLNLKSLIVEADAOORDHAMEGEGERRRAGALLLSDDLSTL-----LRT 245
      ||| : ||| ||| ||| |||
Oy 75 AVVSLGSASASIAOEPAGE-----LVAEEDDOPSELPQTEFSQOPAPPLNLRPRR 128
      ||| : ||| ||| ||| |||
Db 246 ARGVAKQPARISEEEREVERKMLALASALECTIDPASHQALRELAQYA-----VEFQW 299
      ||| : ||| ||| ||| |||
Oy 129 SAPKGRKTRARRAIAAHYEVHPRPQDGAQAGVDGTVSGMEEARINSSSPLYNR--QIG 186
      ||| : ||| ||| ||| |||
Db 300 SNDQ-RYLLTRCISVLLKAVN-----AEKGMRAVASGEVEGRVSGAGTSLMHRDLOMA 351
      ||| : ||| ||| ||| |||
Oy 187 EPIYIRAL-----YLY-----COVHDEBKAVYKLDLVDGVLA 223
      ||| : ||| ||| ||| |||
Db 352 LFYGRSLALLGLSYVIITAMPAPASGAMLLAAVCSLEFRNDNAVAIGSLRGTYA 411
      ||| : ||| ||| ||| |||
Oy 224 LRCEEFSATASSLGPOLRICOVSG--LLALRPSSLRIRTLPMWHLKAAPFLTYEGL 280
      ||| : ||| ||| ||| |||
Db 412 I-----PAAMLVSQMLPQWNGFPLLCLAGVPLFFATLGMAYVVTAGTATSFAL 461
      ||| : ||| ||| ||| |||

RESULT 15
H83044
2,4-dienoyl-CoA reductase FadH2 PA4814 [imported] - Pseudomonas aeruginosa (strain PA
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence-revision 15-Sep-2000 #text-change 31-Dec-2000
C:Accession: H83044
R:Stover, C.K.; Yam, X.Q.; Errvin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; L
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kaser, A.; Larbig, K.; L
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83044
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-681 <STO>
A:Cross-references: GB:AE004894; GB:AE004091; MID:99591076; PIDN:AAG08199.1; GSPDB:GN
C:Genetics:
C:Experimental source: strain PA01
C:Superfamily: Methylotrophilus methylotrophus W3A1 trimethylamine dehydrogenase

Query Match 6.3%; Score 91; DB 2; Length 661;
Best Local Similarity 26.3%; Pred. No. 17;
Matches 79; Conservative 31; Mismatches 86; Indels 104; Gaps 20;

Oy 6 FEISARRLPRLPRSGISRGAGVROAOPAP-----MAARRSORRRG----- 47
      ||| : ||| ||| ||| |||
Db 442 FRVRLERLGVDLRLGHR-----VROGELDGGOFDDVYVATGIDPRRPRIDIGPTVLSYVD 497
      ||| : ||| ||| ||| |||
Oy 48 --RGEP-GTALLVPLAAGLALACGLLLAAVSLGSFASIAOEPAPOEALVAEEDDPS 104
      ||| : ||| ||| ||| |||
Db 498 VLRGAPGARVAIVAGAGIGFDVA--AFVLAAPSG-----OPRALGEMIALMGVDLD 548
      ||| : ||| ||| ||| |||

```

```

Oy 105 ELNQTTESSODRPAEFLNRLVPRRSAAKGGKTTARRRILAAHYEVNRPQDDGAQADVD 164
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 549 - -NSOPGGLRPPAP-----TRPAQV-----WLLORRKPAGQL---GK 583
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Oy 165 VSGMEEARINSSPLRYN--RQIGFIVTRAGLYLYCQVHPBGAKVYLKLD-----LL 217
      |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 584 TSGWVH-----RAHLRHNAAVMLG-----GVETL-----KIDE-KQLIRVUGVERWLE 626
      |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Oy 218 VDGVL-----ALRCLEFESATASSLGPOLRL-----COVSGLL-----ALRPGSSLRIR 262
      |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 627 VDNVVICAGDEPRLDELQDROTSES-----LRPILIGAGANVAGELDQAKRAIRRGAMLAIR 680
      |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:

```

Search completed: March 31, 2003, 07:59:21  
Job time : 34.4558 secs



STIC-Biotech/Ch mLib

89 055

From: Schnizer, Richard  
Sent: Friday, March 14, 2003 6:45 AM  
To: STIC-Biotech/ChemLib  
Subject: 09/245,198

RECEIVED  
MAR 14 2003  
STIC

Please search the commercial and interference databases for SEQ ID NOS: 1-4 of 09/245,198.

Please perform an oligo search on SEQ ID NO:4.

Please also search for nucleic acids that can encode the amino acid sequences of SEQ ID NOS:2 or 4.

Thank you-

Richard Schnizer, Ph.D.  
Patent Examiner  
Art Unit 1635  
CM1 12E17  
703-306-5441  
Mail Box CM1 11E12

Point of Contact:  
Thomas G. Larson, Ph.D.  
703-308-7309  
CM1, Rm. 6 B 01

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 3/24  
Date Completed: 4/1  
Searcher Prep/Review: 15  
Clerical: \_\_\_\_\_  
Online time: 20

TYPE OF SEARCH: 4  
NA Sequences: \_\_\_\_\_  
AA Sequences: 3  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: ABSSOT  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**